Business Report - 8

PG Program in Data Science and Business Analytics

submitted by

Sangram Keshari Patro

BATCH:PGPDSBA.O.AUG24.B

Great Learning

POWER AHEAD

Contents

List of Figures

1 Objective

The primary objective of this project is to analyze and forecast wine sales trends for the 20th century based on historical data provided by ABC Estate Wines. We aim to equip ABC Estate Wines with the necessary insights and foresight to enhance sales performance, capitalize on emerging market opportunities, and maintain a competitive edge in the wine industry.

2 Data Description

The data provided is of sales of two types of wine Rose and Sparkling from the period of 1980 to 1995.

3 Data Overview

3.1 Importing necessary libraries and the dataset

Datasets for both types of wine are printed. Both have 187 rows & 2 columns.

3.2 Structure and type of data

Data is explored further. The dataset is free from duplicate rows and contains some null values.

```
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   YearMonth  187 non-null    datetime64[ns]
 1   Rose       185 non-null    float64
dtypes: datetime64[ns](1), float64(1)
memory usage: 3.1 KB
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 187 entries, 0 to 186
```

```
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   YearMonth  187 non-null    datetime64[ns]
 1   Sparkling  187 non-null    int64
dtypes: datetime64[ns](1), int64(1)
memory usage: 3.1 KB
```

Figure 1: Table depicting the datatype for both types of wine.

3.3 Statistical summary

	YearMonth	Rose
count	187	185.000000
mean	1987-10-01 08:51:20.213903744	90.394595
min	1980-01-01 00:00:00	28.000000
25%	1983-11-16 00:00:00	63.000000
50%	1987-10-01 00:00:00	86.000000
75%	1991-08-16 12:00:00	112.000000
max	1995-07-01 00:00:00	267.000000
std	NaN	39.175344

	YearMonth	Sparkling
count	187	187.000000
mean	1987-10-01 08:51:20.213903744	2402.417112
min	1980-01-01 00:00:00	1070.000000
25%	1983-11-16 00:00:00	1605.000000
50%	1987-10-01 00:00:00	1874.000000
75%	1991-08-16 12:00:00	2549.000000
max	1995-07-01 00:00:00	7242.000000
std	NaN	1295.111540

Figure 2: Statistical summary of the data

4 Exploratory Data Analysis

4.1 Univariate Analysis

4.1.1 'Rose'

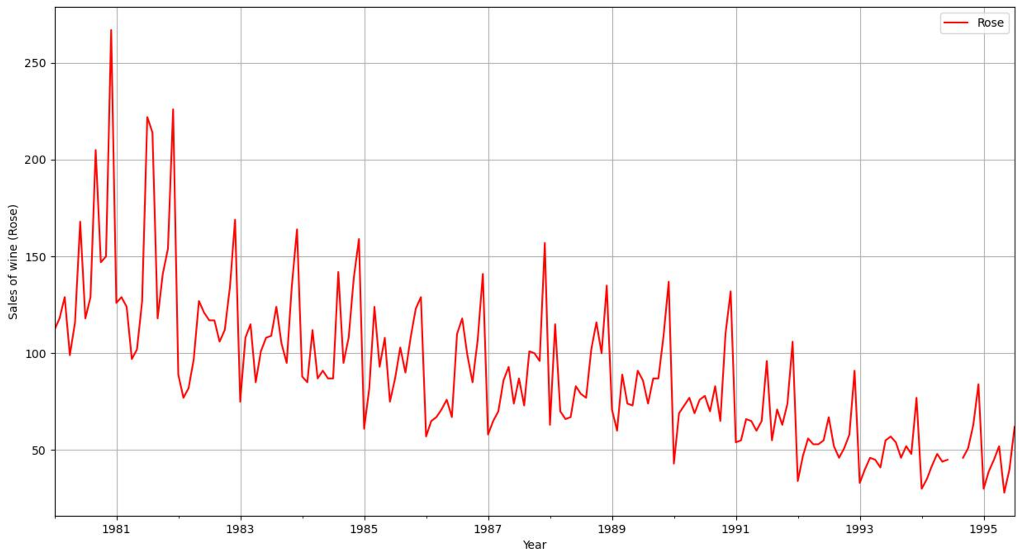


Figure 3: Timeseries plot of **'Rose'** wine sales

Observations

- **Sales Trend**: There is a noticeable overall downward trend in the sales of rosé wine from 1980 to 1995.
- **Seasonality**: Spikes in sales occur regularly, indicating a possible seasonal pattern.
- **Volatility**: Early years (1980–1983) show high fluctuation in sales, which stabilizes over time.
- **Peak Period**: The highest sales were recorded around 1980–1982, after which a gradual decline is observed.

Business Recommendations

- **Seasonal Marketing**: Utilize peak seasons to promote rosé wine through targeted campaigns and discounts.
- **Revitalization Strategy**: Investigate causes of declining demand and consider rebranding or new packaging to rejuvenate consumer interest.
- **Diversify Portfolio**: Consider expanding into other wine categories or beverages to offset the decline in rosé sales.
- **Customer Analysis**: Conduct market research to understand changing preferences and tailor offerings accordingly.

4.1.2 'Sparkling'

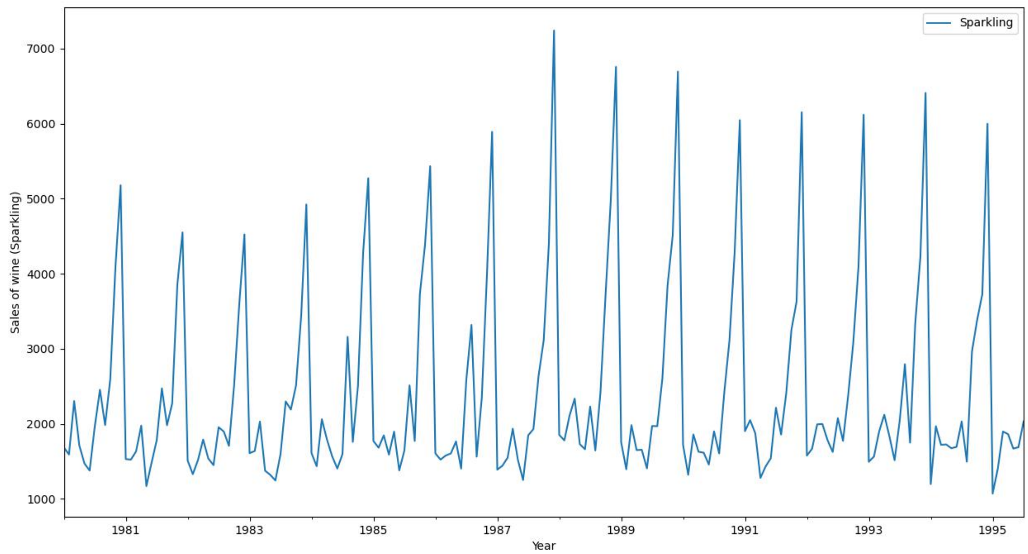


Figure 4: Timeseries plot of **'Sparkling'** wine sales

Observations

- **Sales Trend**: Overall, there is a clear upward trend in sparkling wine sales over the years.
- **Seasonality**: Sales show consistent yearly spikes, likely during festive seasons.
- **Volatility**: Peaks are sharp and regular, indicating predictable consumer behavior patterns.
- **Growth**: From 1980 to 1995, both the baseline and peak sales values have increased steadily.

Business Recommendations

- **Seasonal Promotions**: Intensify marketing during known high-demand months (e.g., holidays) to maximize sales.
- **Capacity Planning**: Scale production and distribution in advance of seasonal peaks.
- **Product Positioning**: Leverage the premium and celebratory image of sparkling wine to appeal to aspirational consumers.
- **International Markets**: Given strong domestic growth, explore export opportunities during peak global festive seasons.

4.2 Bivariate Analysis

4.2.1 Rose sales vs Year

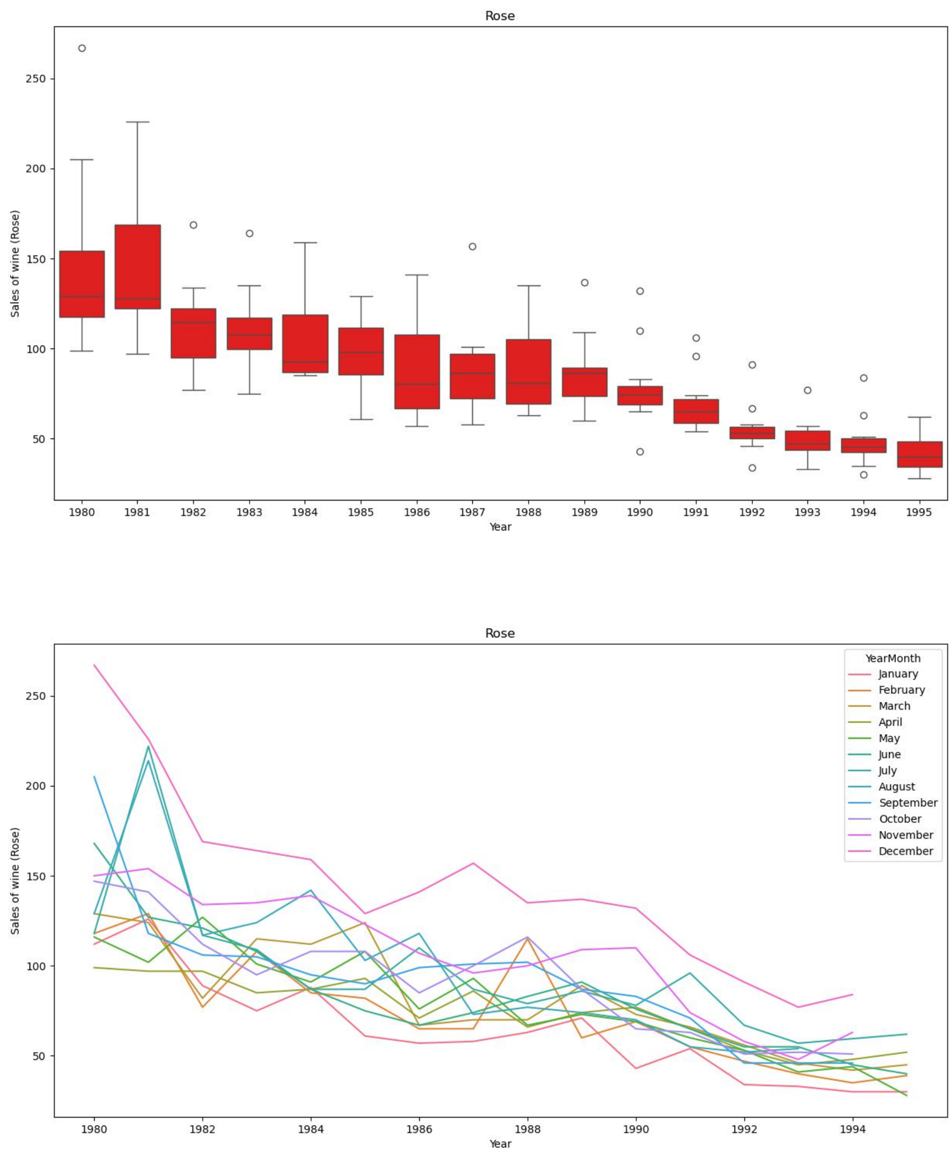


Figure 5: Rose sales vs Year

Observations

- **Box Plot**: Sales of rosé wine declined over the years. Variability and median decreased, with more outliers in early years indicating occasional high sales.
- **Monthly Trends**: November and December show consistently higher sales, suggesting seasonality. All months show a general downward trend.

Business Recommendations

- **Seasonal Targeting**: Focus promotions around November–December due to peak sales in these months.
- **Revival Strategy**: Revamp marketing for rosé wine to counter steady decline over time.
- **Outlier Utilization**: Analyze high-sales months in early years to replicate successful strategies.

4.2.2 Sparkling sales vs Year

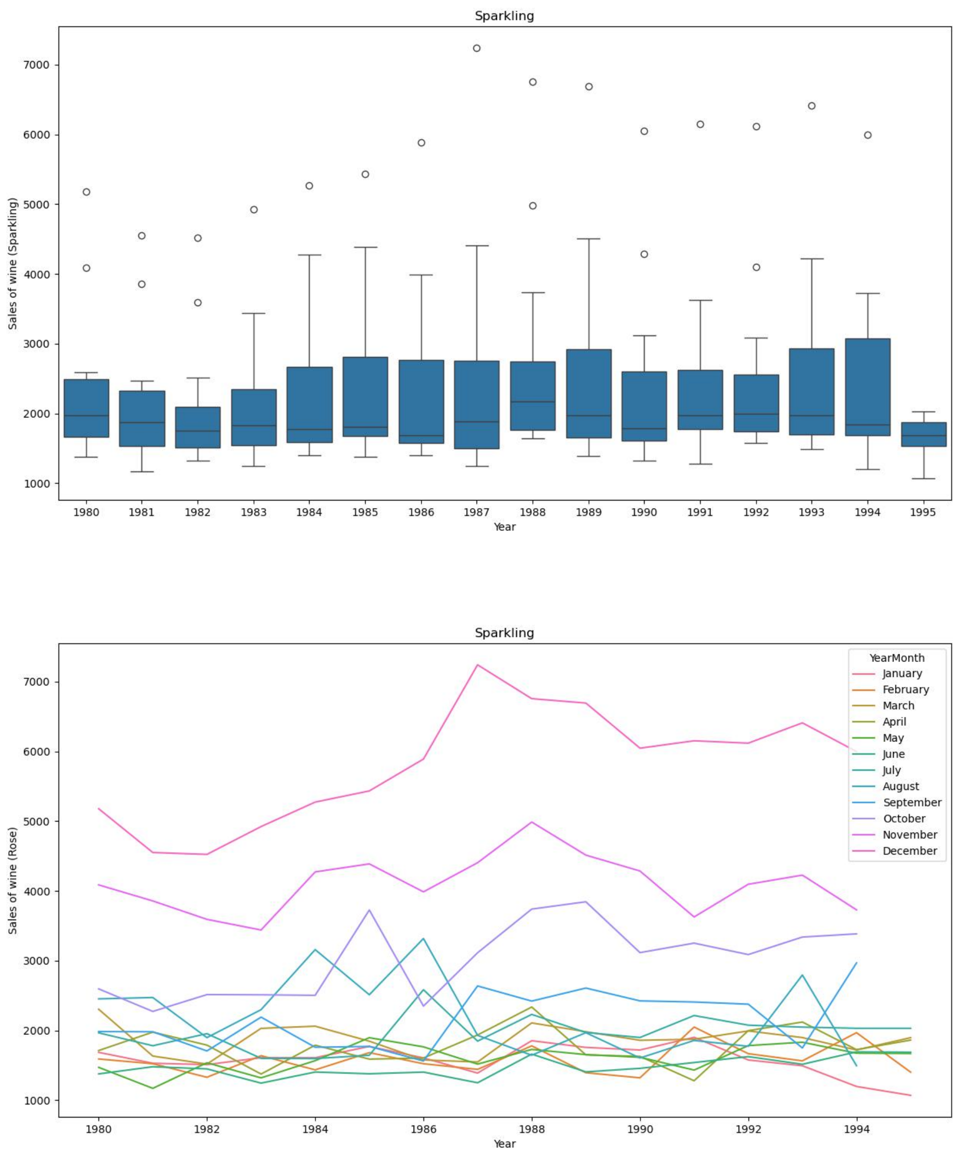


Figure 6: Sparkling sales vs Year

Observations

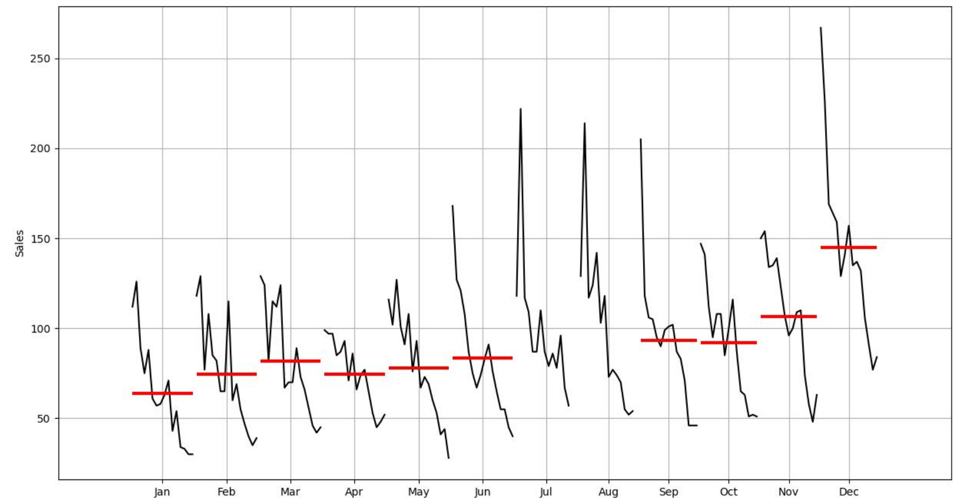
- **Box Plot (Sparkling Wine)**: Sales increased slightly from 1980 to 1995. Distribution remained steady, but outliers rose after 1985, showing sporadic sales spikes.

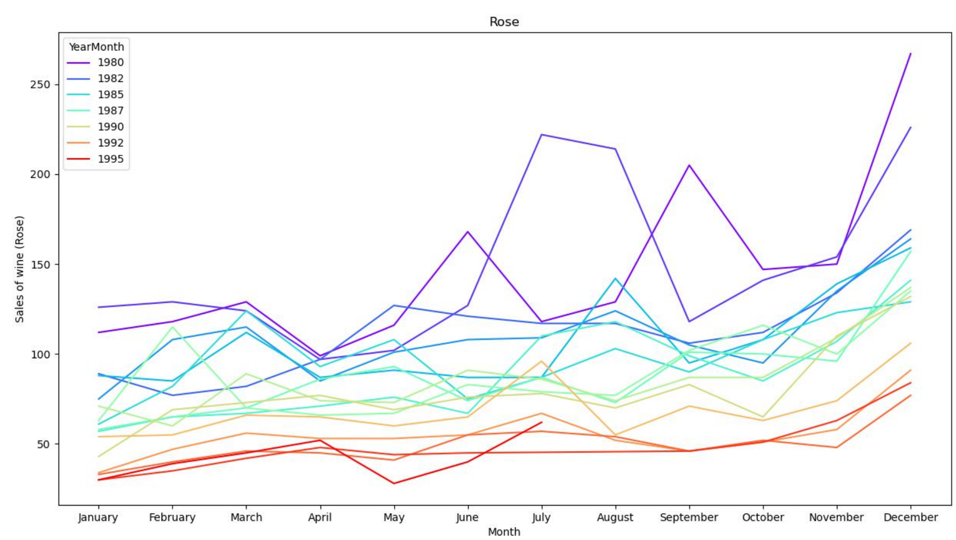
- **Monthly Trends (Sparkling Wine)**: Clear seasonality with high sales in November and December. Other months showed stable but lower sales over time.

Business Recommendations

- **Seasonal Focus**: Intensify campaigns in November–December to leverage strong seasonal demand.
- **Peak Analysis**: Investigate reasons behind outlier months with high sales to replicate success.
- **Stable Base**: Maintain consistent supply across months while optimizing for year-end spikes.

4.2.3 Rose sales vs Month





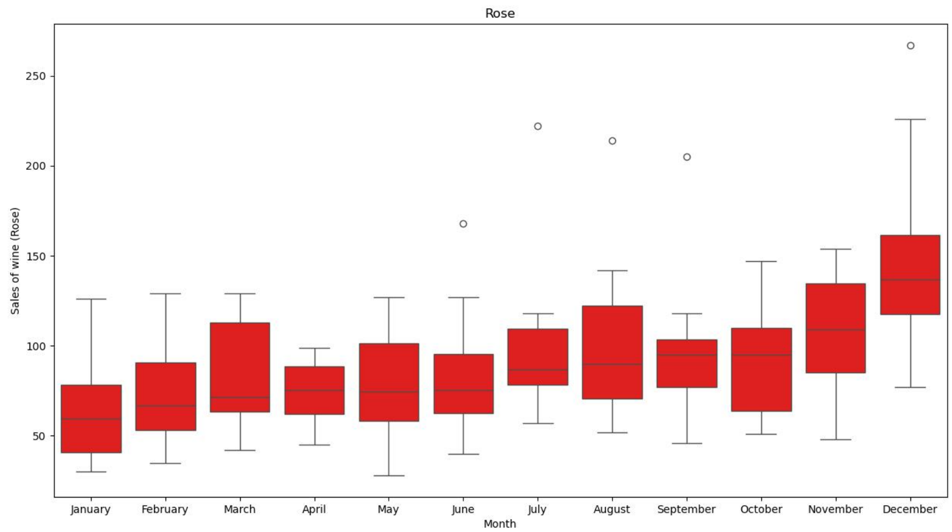


Figure 7: Rose sales vs Month

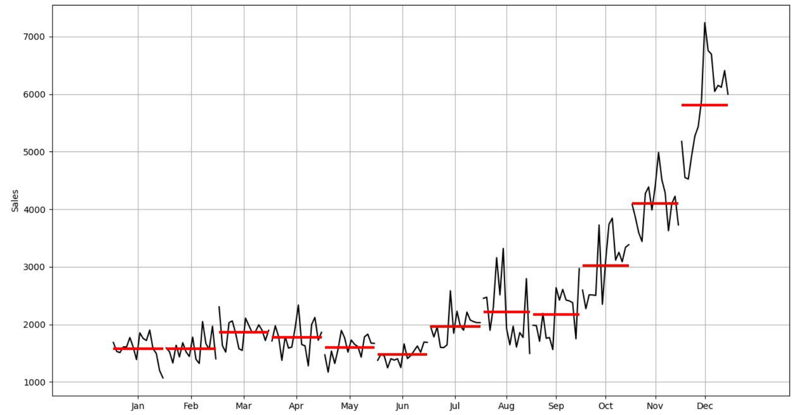
Observations

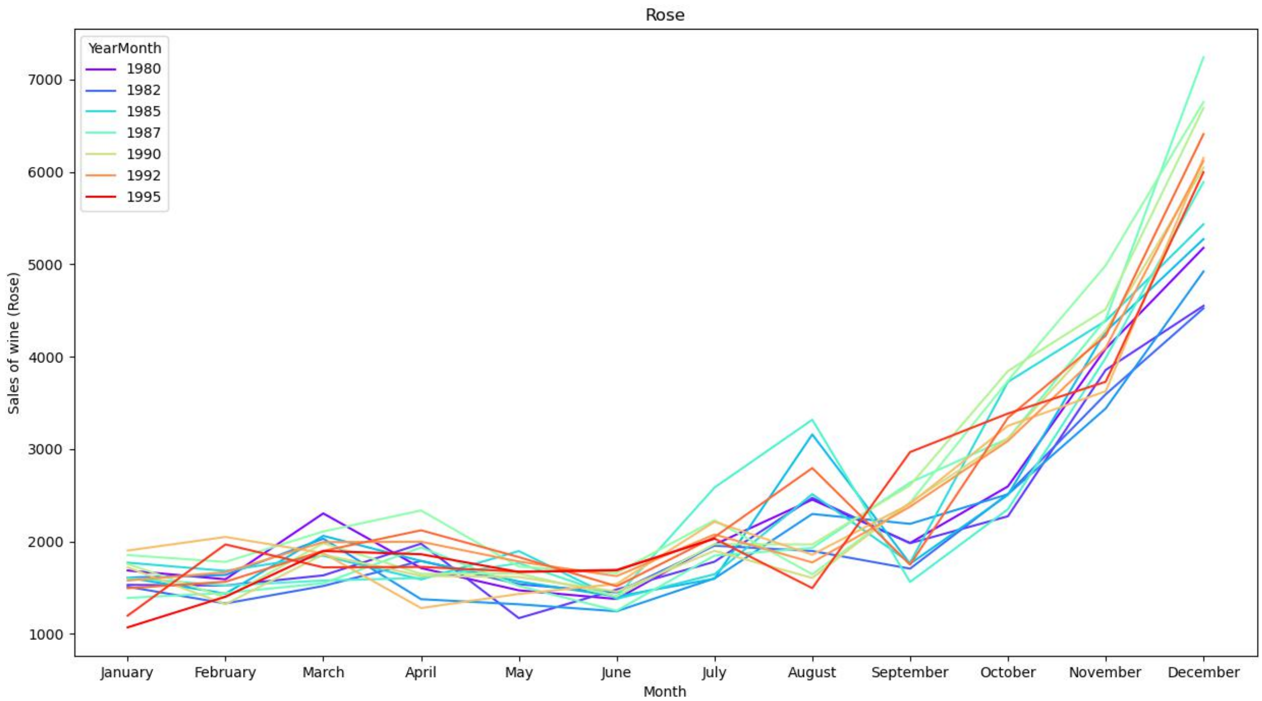
- **Trend Plot**: Clear upward trend from January to December; highest spikes in November–December.
- **Yearly Comparison**: Sales peaked in the 1980s; gradual decline post-1990. Seasonality remains strong across years.

Business Recommendations

- **Seasonal Focus**: Boost marketing in late Q4 to leverage peak sales.
- **Sales Recovery**: Study 1980s peak years to extract replicable growth strategies.
- **Product Strategy**: Consider bundling or festive packaging for holiday months.

4.2.4 Sparkling sales vs Month





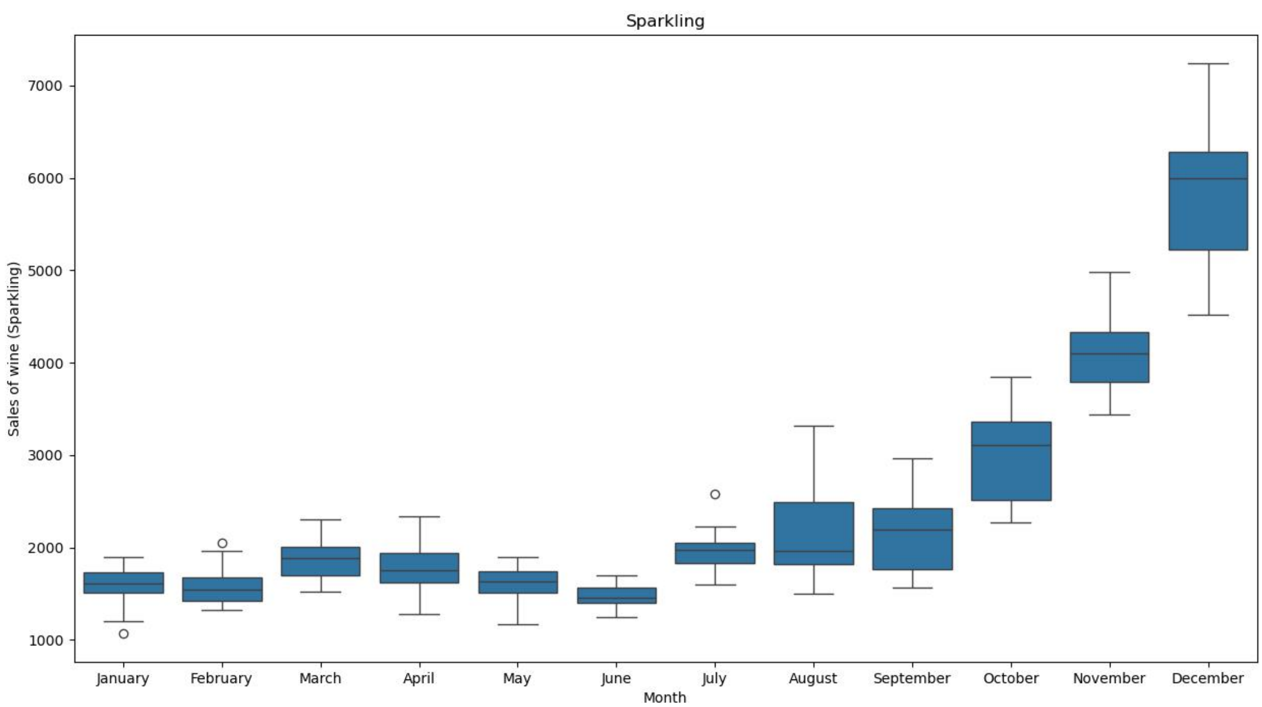


Figure 8: Sparkling sales vs Month

Insights

- **Trend Plot**: Sales remain flat in the first half, then rise sharply from July to December, with steep peaks in November–December.
- **Yearly Comparison**: Consistent year-end spikes across all years; 1995 shows strongest finish. Overall upward trend post-June.

Business Recommendations

- **Holiday Strategy**: Prioritize inventory and campaigns for Q4 to align with demand surges.
- **Growth Momentum**: Scale up operations mid-year to prepare for strong second-half sales.
- **Year-End Promotions**: Launch targeted festive offers and high-value bundles to boost conversions.

4.3 Plotting the Empirical Cumulative Distribution

This particular graph tells us what percentage of data points refer to what number of Sales.

4.3.1 Rose sales

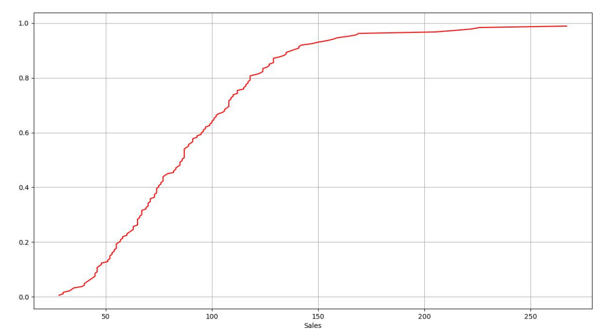


Figure 9: **ECDF plot of Rose sales**

4.3.2 Sparkling sales

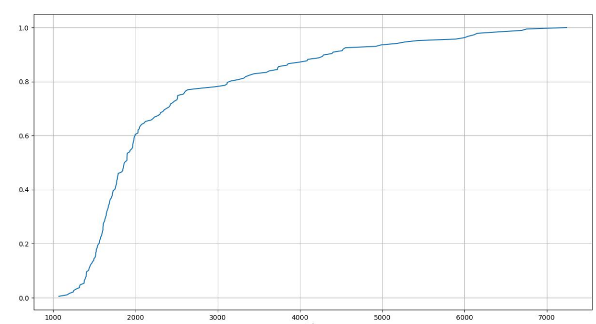


Figure 10: **ECDF plot of Sparkling sales**

4.4 Plotting the average RetailSales per month and the month on month percentage change of RetailSales.

4.4.1 Rose sales

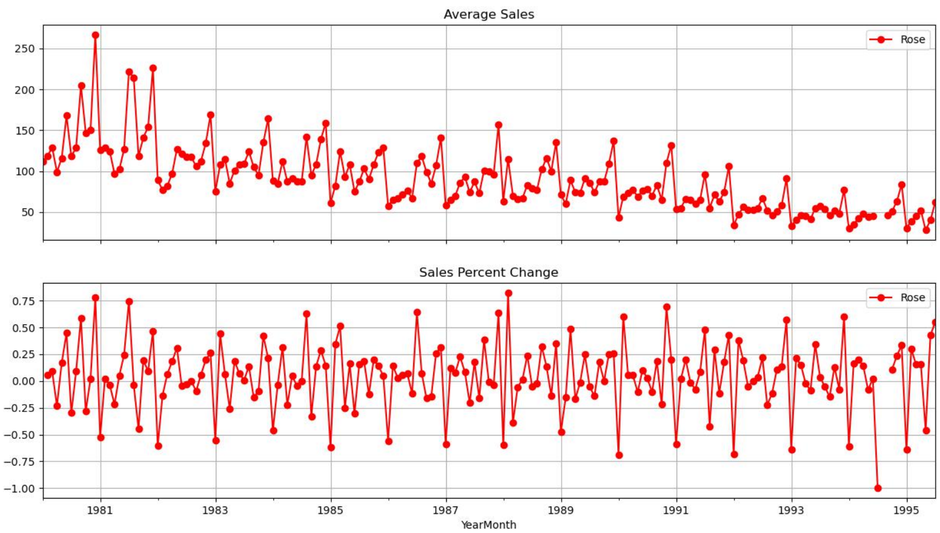


Figure 11: **Rose sales**

4.4.2 Sparkling sales

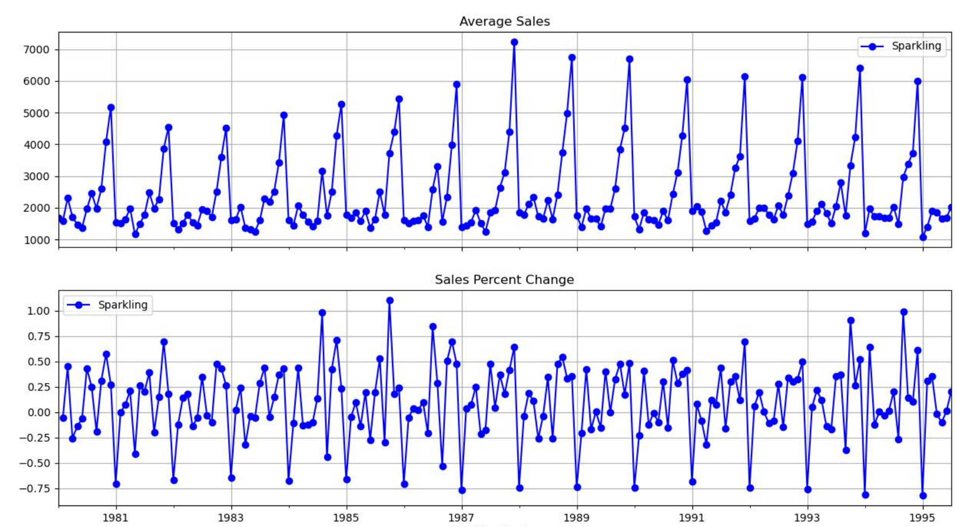


Figure 12: **Sparkling sales**

5 Data Pre-processing

I have forward filled the 2 null values encountered in Rose dataset and no null value is found is Sparkling dataset.

	Rose
YearMonth	
1994-07-01	NaN
1994-08-01	NaN

After removing null values data is decomposed into trend,seasonality and residuals and then visualized by plotting.

Perform Decomposition

5.1 Additive Model

5.1.1 Rose sales

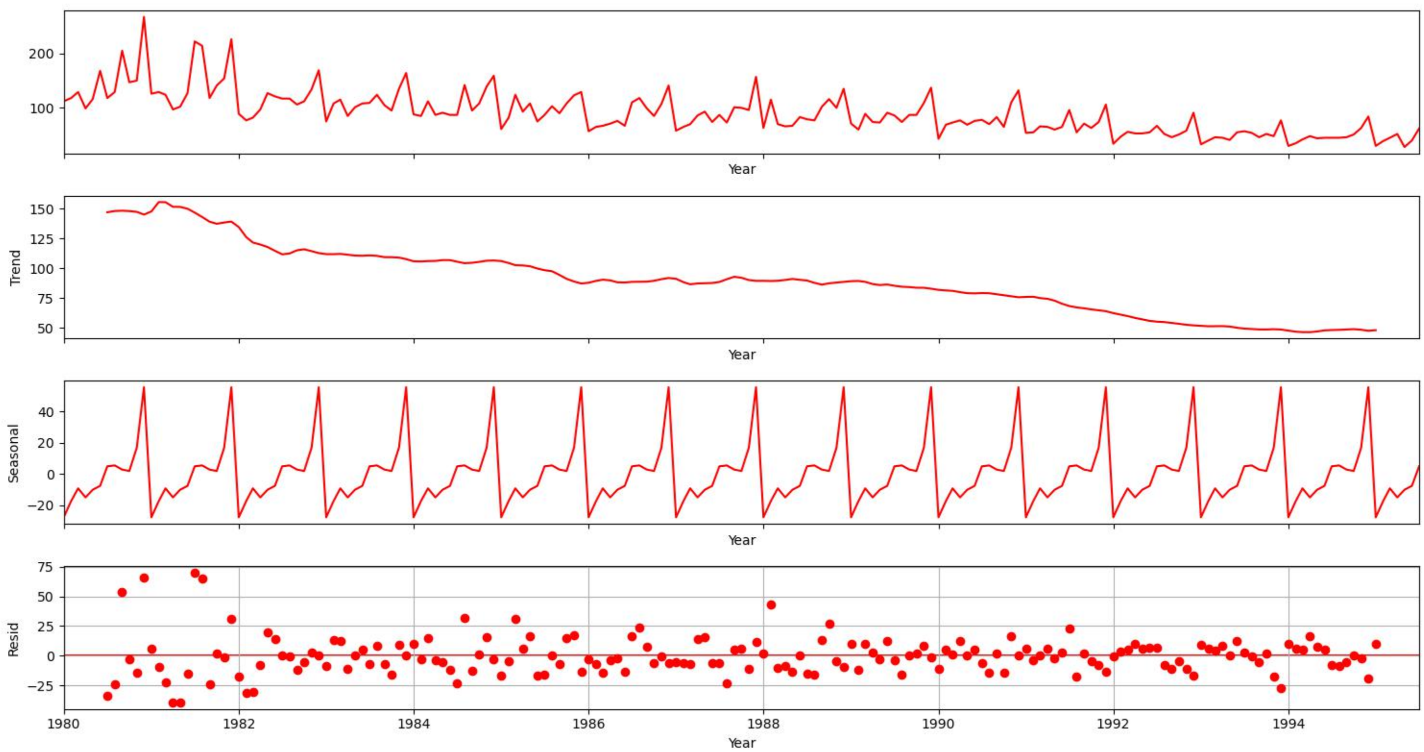


Figure 13: **Rose sales**(Additive Model)

Insights

- **Trend Plot**: Long-term decline observed in overall values post-1981, indicating a steady downward trend.
- **Seasonality**: Clear recurring peaks annually, with strong cyclical behavior suggesting seasonal effects.
- **Residuals**: Randomly scattered, indicating minimal pattern left; model captures most structure.

Business Recommendations

- **Seasonal Planning**: Leverage predictable seasonal peaks for campaign timing and resource allocation.
- **Decline Strategy**: Investigate causes for long-term decline and strategize for revival.
- **Model Reliability**: Proceed with forecasting, as residuals show no significant patterns.

5.1.2 Sparkling sales

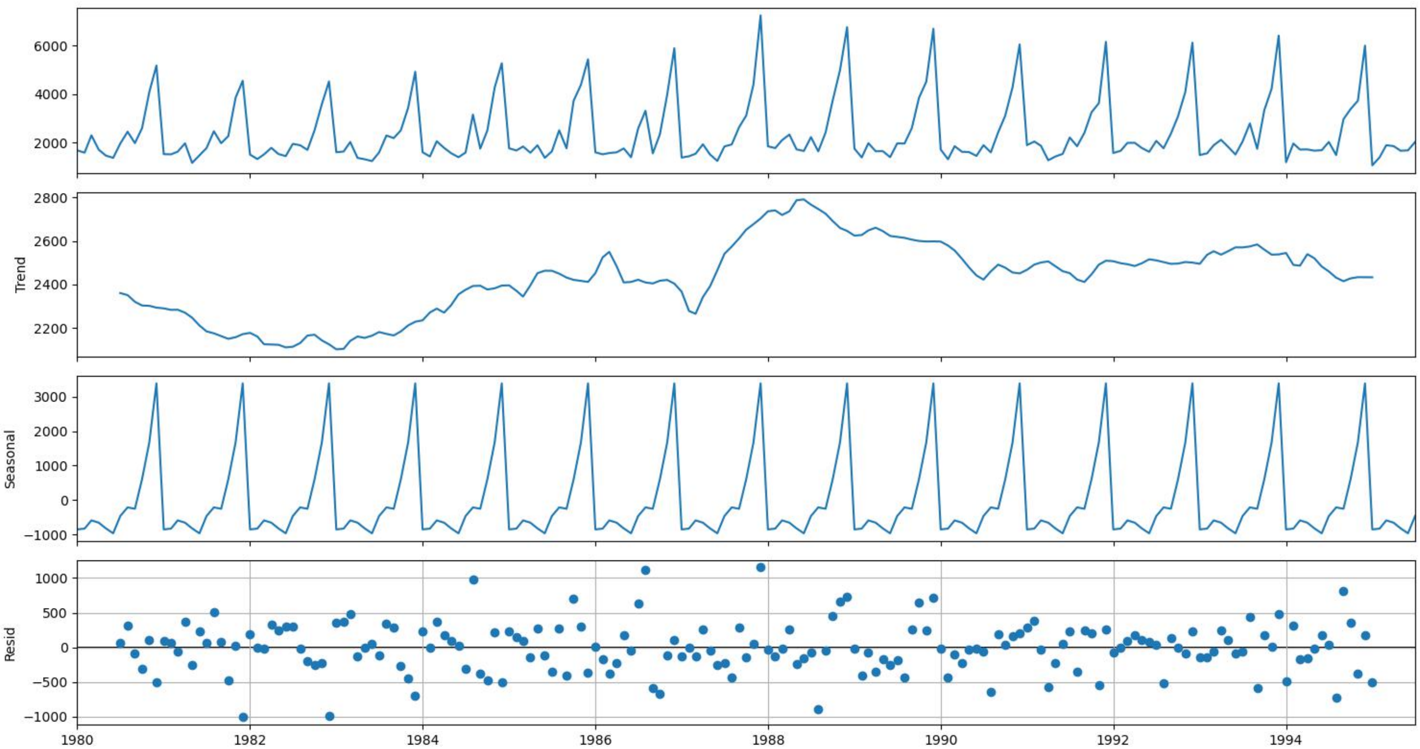


Figure 14: **Sparkling sales(**Additive Model**)**

Insights

- **Trend Plot**: Mild dip till 1983, followed by a steady rise peaking around 1989–1990, then slight decline.
- **Seasonality**: Strong, consistent seasonal spikes each year, showing high demand fluctuations.
- **Residuals**: Randomly scattered around zero with minor variance; model fits the data well.

Business Recommendations

- **Peak Timing**: Focus marketing and logistics around predictable annual spikes.
- **Capacity Planning**: Align operations with the rising trend observed mid-80s to early-90s.
- **Monitoring Needed**: Investigate post-1990 plateau or decline to inform future strategy.

5.2 Multiplicative Model

5.2.1 Rose sales

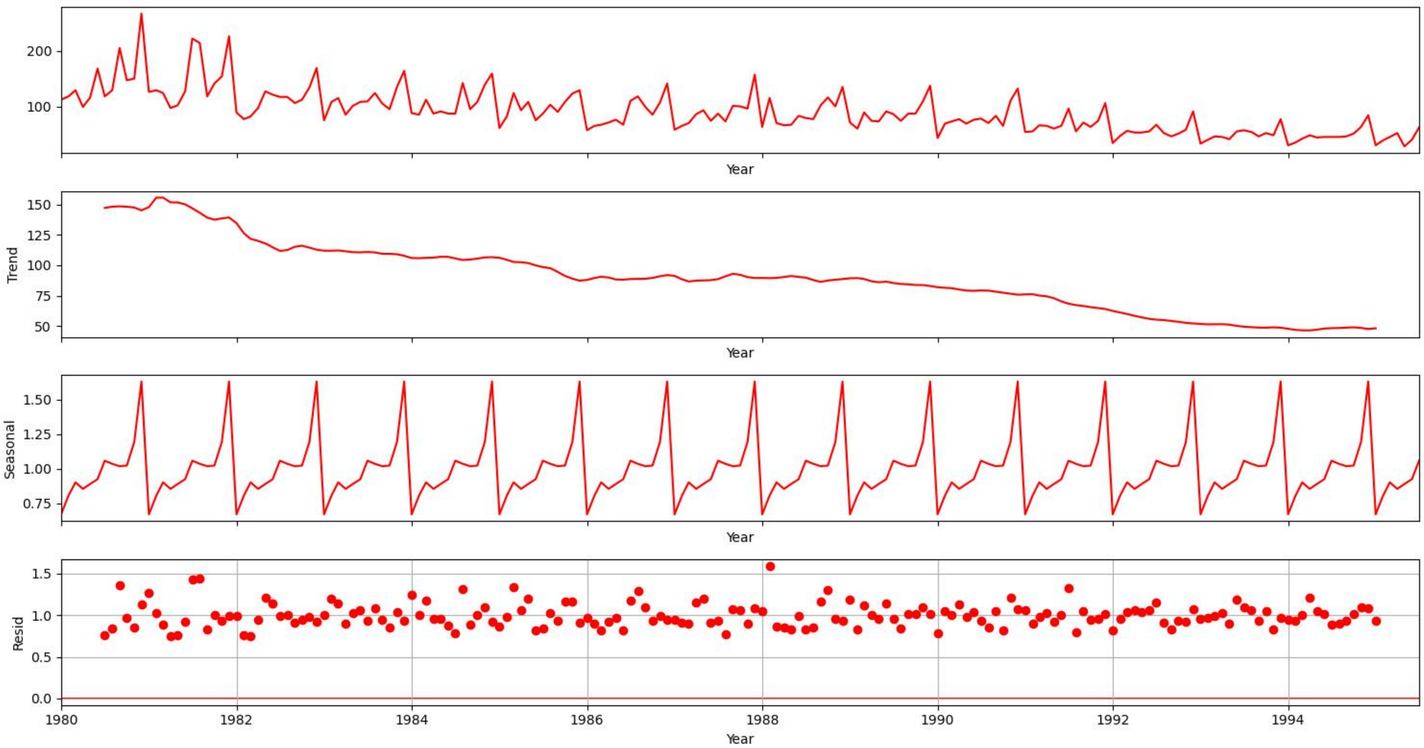


Figure 15: **Rose sales**(Multiplicative Model)

Insights

- **Trend Plot**: Noticeable decline throughout the period, indicating weakening growth over time.
- **Seasonality**: Clear yearly patterns with relative seasonal effects remaining proportional over time.
- **Residuals**: Fluctuations appear stable and centered, suggesting a good model fit.

Business Recommendations

- **Stabilize Decline**: Identify and mitigate drivers of downward trend.
- **Seasonal Focus**: Seasonal proportionality suggests timing remains key – optimize around peak months.
- **Model Validity**: Continue using multiplicative models for accurate forecasting and planning.

5.2.2 Sparkling sales

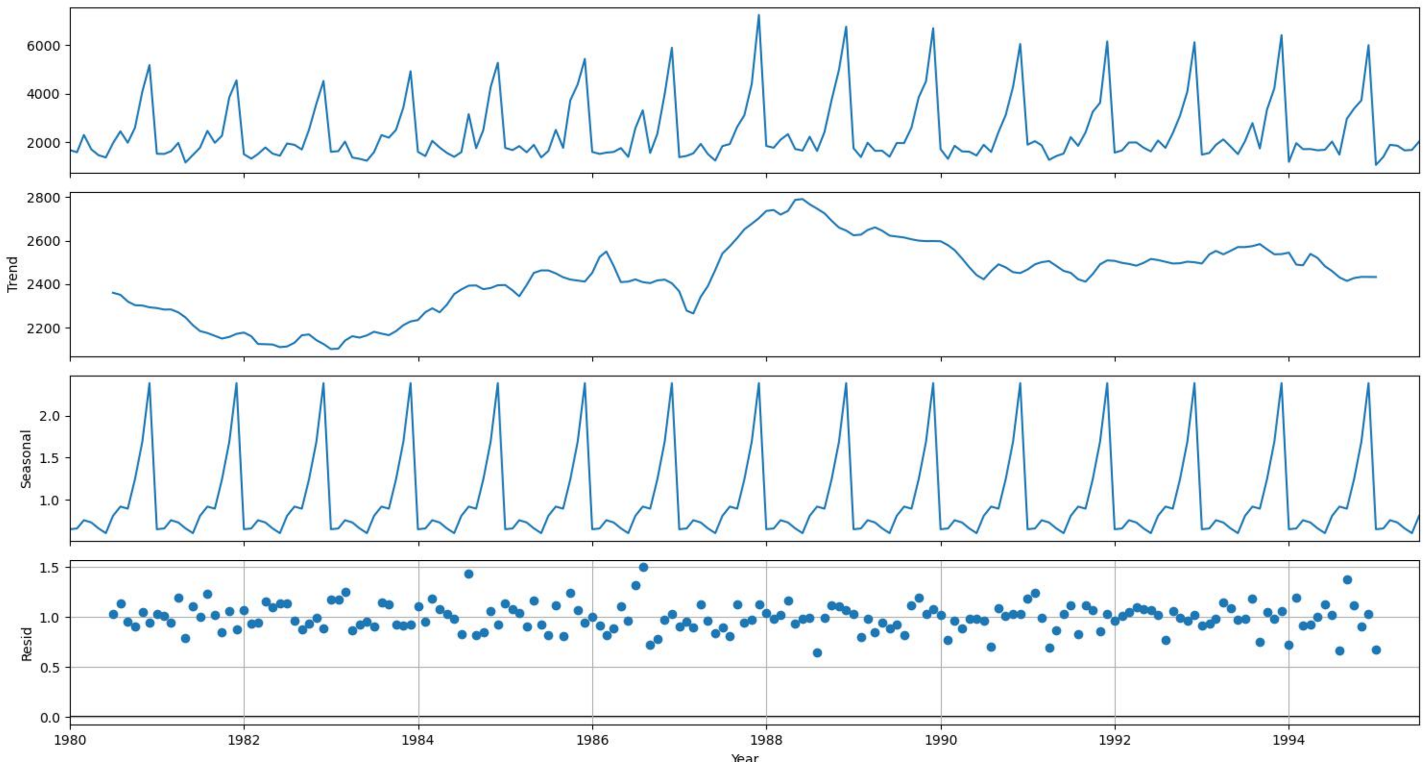


Figure 16: **Sparkling sales**(Multiplicative Model)

Insights

- **Trend Plot**: Gradual decline in long-term trend, indicating reduced popularity or demand.
- **Seasonality**: Strong multiplicative seasonality – peaks maintain proportional relationship to trend.
- **Residuals**: Stable and random around 1, supporting model appropriateness.

Business Recommendations

- **Address Downtrend**: Explore marketing or innovation strategies to reverse trend.
- **Leverage Seasonality**: Plan production and promotions around predictable seasonal peaks.
- **Model Reliability**: Multiplicative model effectively captures structure – use for forecasting.

5.3 Model Comparison Summary

- **Rose Data:** Multiplicative model shows clearer seasonal variation with stable residuals. Best fit.
- **Sparkling Data:** Multiplicative model better captures proportional seasonality with a declining trend. Best fit.

Conclusion

Multiplicative models are more suitable for both Rose and Sparkling datasets due to their ability to model varying seasonal effects relative to trend levels. Then the data is split into train data and test data.

6 Build forecasting models

6.1 Linear Regression

6.1.1 Model Building

Linear regression model is built on train data and fitted on test data to get the following plot.

Rose Sales

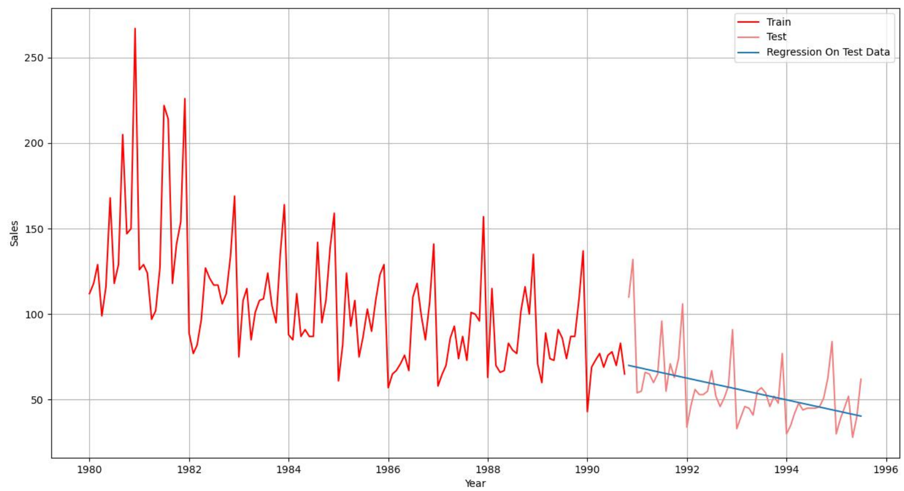


Figure 17: **Rose sales**(Linear Regression)

Sparking Sales

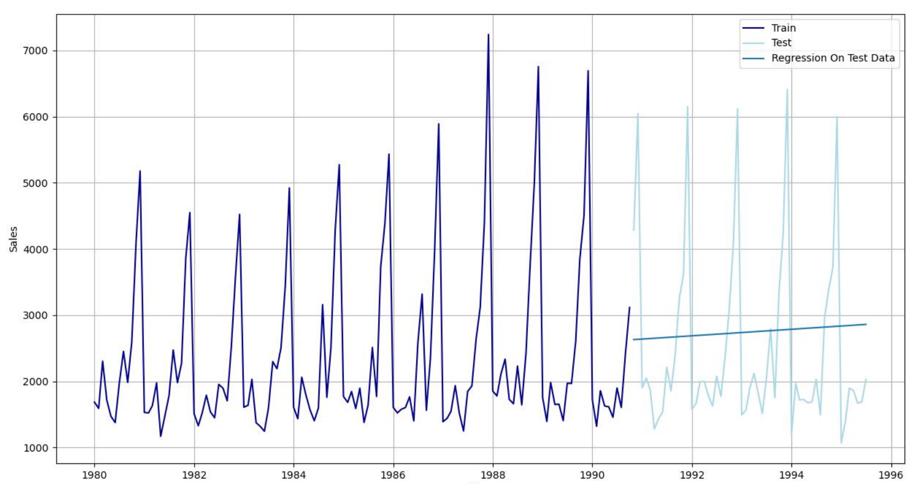


Figure 18: **Sparkling sales**(Linear Regression)

6.1.2 Model Evaluation

The RMSE for the forecast is given below.

	Test RMSE(Rose)	Test RMSE(Sparkling)
Linear_Regression	17.356924	1392.438305

6.2 Simple Average

6.2.1 Model Building

Simple Average model is built on train data and fitted on test data to get the following plot.

Rose Sales

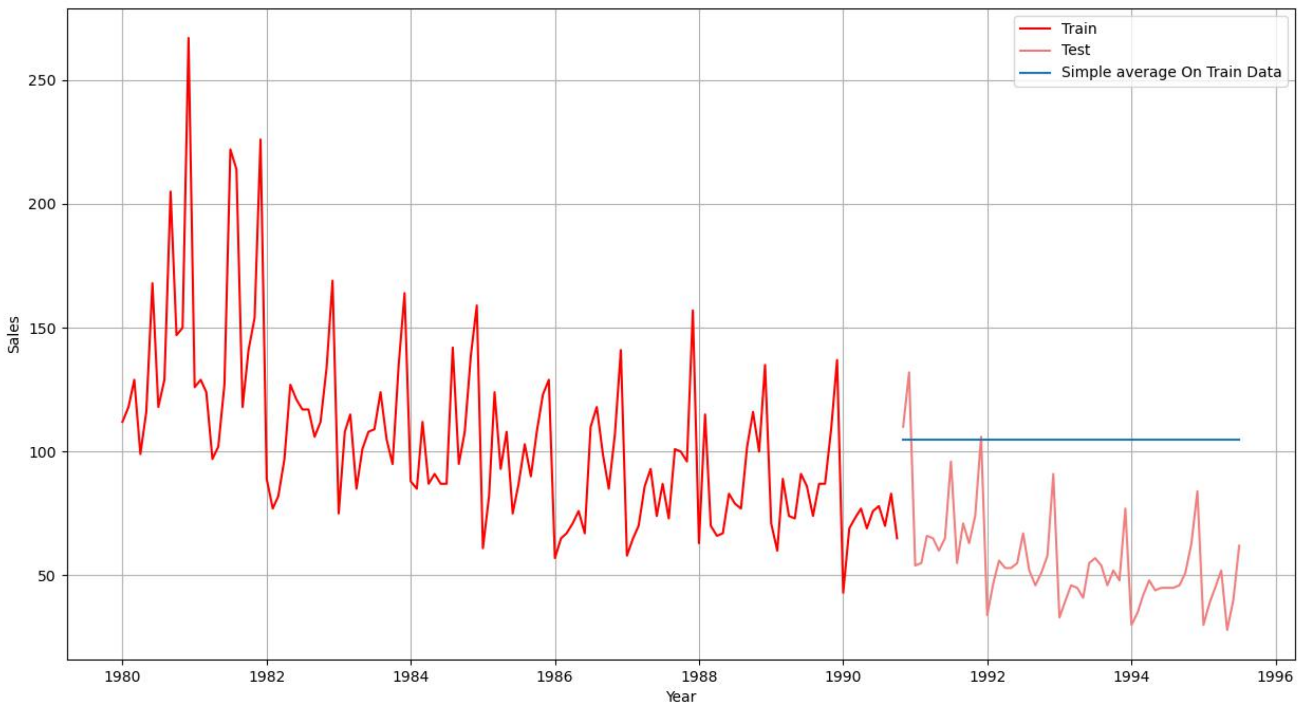


Figure 19: **Rose sales(**Simple Average**)**

Sparking Sales

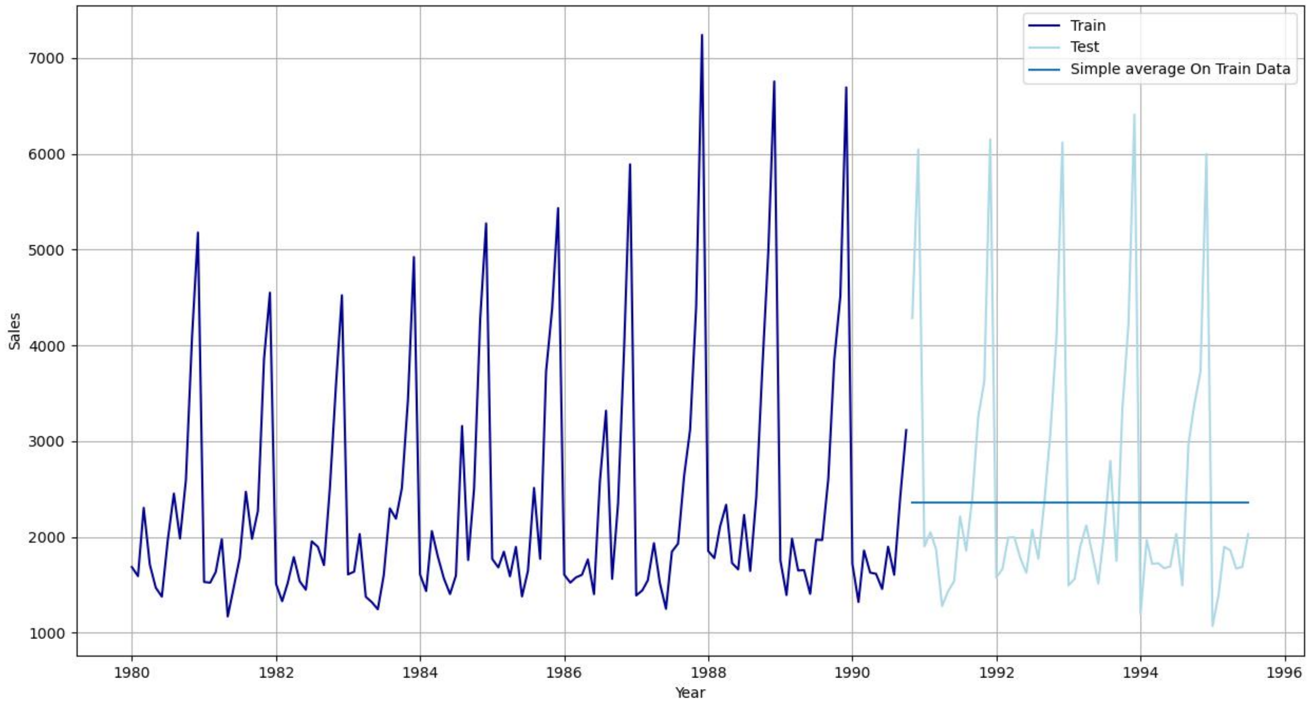


Figure 20: **Sparkling sales(**Simple Average**)**

6.2.2 Model Evaluation

The RMSE for the forecast is given below.

	Test RMSE(Rose)	Test RMSE(Sparkling)
Simple_Average	52.431977	1368.746717

6.3 Moving Average

6.3.1 Model Building

Moving Average model is built on the whole data to get the following plot. The model is built with different window sizes i.e. different point moving averages and best fit is found to be 2-point moving average by looking at the least RMSE.

Rose Sales

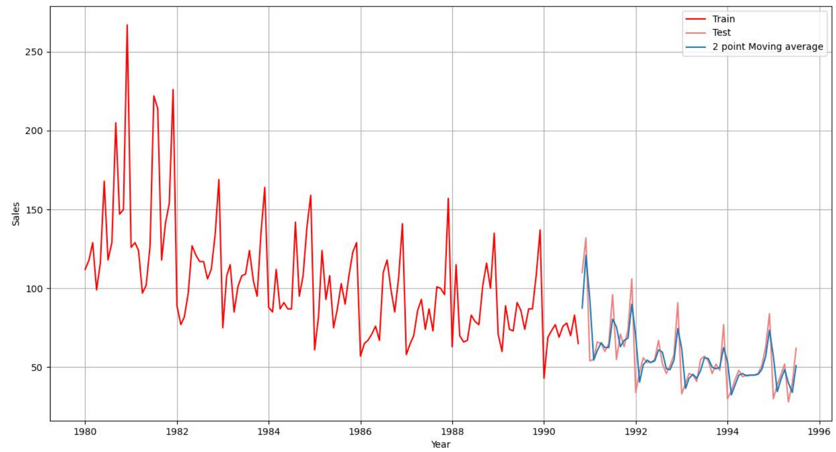


Figure 21: **Rose sales(**Moving Average**)**

Sparking Sales

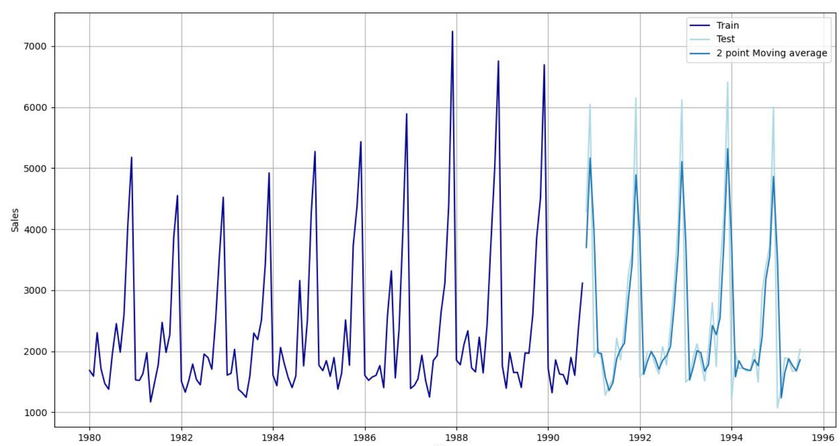


Figure 22: **Sparkling sales(**Moving Average **)**

6.3.2 Model Evaluation

The RMSE for the forecast is given below.

Dataset	Rose	Sparkling
Window_Size		
2	11.801167	811.178937
3	14.769245	1040.067324
4	15.370676	1184.213295
5	15.647319	1276.589323
6	15.867384	1337.200524
7	16.456661	1393.366184
8	16.341787	1412.787236
9	16.345032	1422.653281

```
Best Performing Models:
```

	Test RMSE(Rose)	Test RMSE(Sparkling)
Moving_Average	11.801167	811.178937

The best models for both dataset are 2-point averages.

6.4 Exponential Models (Single, Double, Triple)

Exponential smoothing methods consist of flattening time series data. Exponential smoothing averages or exponentially weighted moving averages consist of forecast based on previous periods data with exponentially declining influence on the older observations. Exponential smoothing methods consist of special case exponential moving with notation ETS (Error, Trend, Seasonality) where each can be none(N), additive (N), additive damped (Ad), Multiplicative (M) or multiplicative damped (Md). One or more parameters control how fast the weights decay. These parameters have values between 0 and 1.

Single Exponential Model

The simplest of the exponentially smoothing methods is naturally called simple exponential smoothing (SES). This method is suitable for forecasting data with no clear trend or seasonal pattern. In Single ES, the forecast at time (t + 1) is given by Winters,1960

$$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t$$

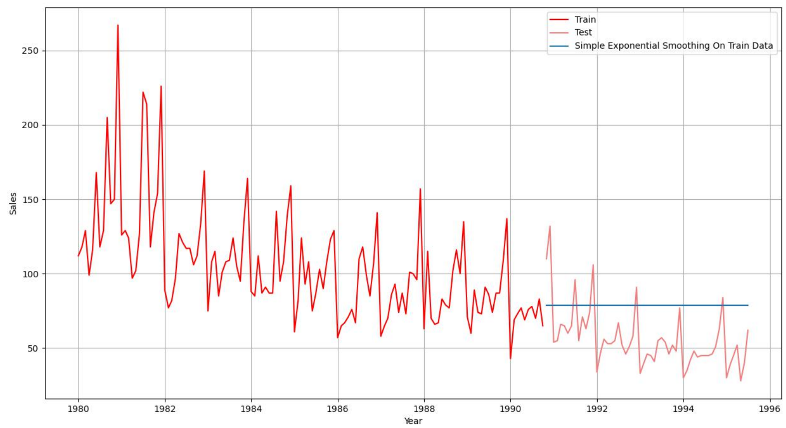
Parameter is called the smoothing constant and its value lies between 0 and 1. Since the model uses only one smoothing constant, it is called Single Exponential Smoothing.

Note: Here, there is both trend and seasonality in the data. So, we should have directly gone for the Triple Exponential Smoothing but Simple Exponential Smoothing and the Double Exponential Smoothing models are built over here to get an idea of how the three types of models compare in this case.

6.4.1 Model Building

Single Exponential Model model is built on train data and fitted on test data to get the following plot.

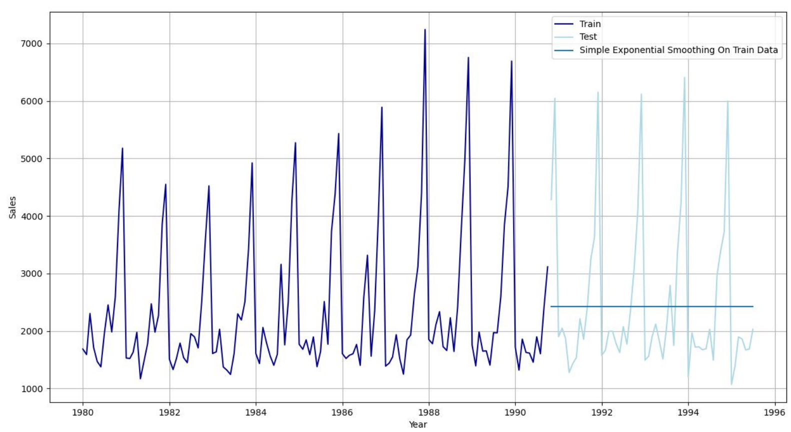
Rose Sales




```
{'smoothing_level': 0.10272100898511088,
 'smoothing_trend': nan,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 134.2627870419287,
 'initial_trend': nan,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

Figure 23: **Rose sales(Single Exponential Model and its parameters)**

Sparking Sales




```
{'smoothing_level': 0.06994041702168571,
 'smoothing_trend': nan,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 2108.034139790532,
 'initial_trend': nan,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

Figure 24: **Sparkling sales(Single Exponential Model and its parameters)**

6.4.2 Model Evaluation

The RMSE for the forecast is given below.

	Test RMSE(Rose)	Test RMSE(Sparkling)
Simple Exponential Smoothing	30.207858	1363.702251

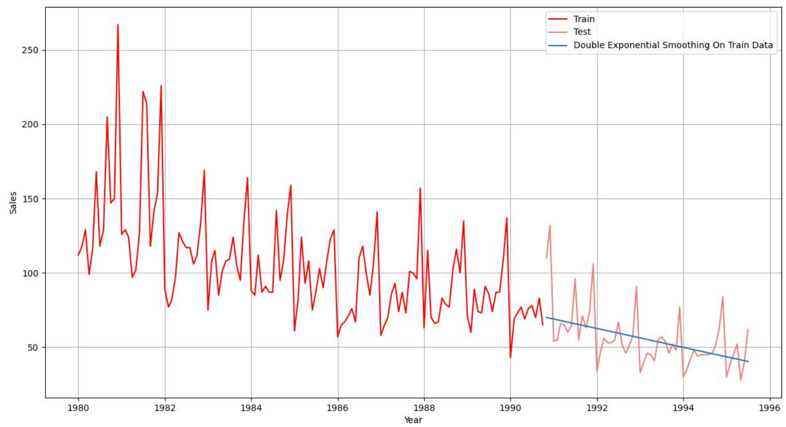
Double Exponential Model

One of the drawbacks of the simple exponential smoothing is that the model does not do well in the presence of the trend. This model is an extension of SES known as Double Exponential model which estimates two smoothing parameters. Applicable when data has Trend but no seasonality. Two separate components are considered: Level and Trend. Level is the local mean. One smoothing parameter α corresponds to the level series A second smoothing parameter β corresponds to the trend series.

6.4.3 Model Building

Double exponential model is built on train data and fitted on test data to get the following plot.

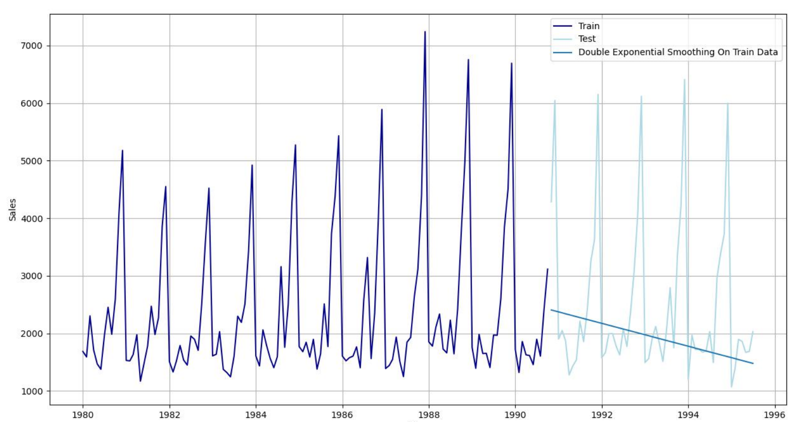
Rose Sales




```
{'smoothing_level': 1.4901161193847656e-08,
 'smoothing_trend': 7.755984441513712e-11,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 139.3527819489728,
 'initial_trend': -0.5291705700335453,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

Figure 25: **Rose sales(Double Exponential Model and its parameters)**

Sparking Sales




```
{'smoothing_level': 0.07614001422051413,
 'smoothing_trend': 0.07614001422051413,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 1505.800070851357,
 'initial_trend': 2.7690564419590373,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

Figure 26: **Sparkling sales(Double Exponential Model and its parameters)**

6.4.4 Model Evaluation

The RMSE for the forecast is given below.

	Test RMSE(Rose)	Test RMSE(Sparkling)
Double Exponential Smoothing	17.356857	1472.253632

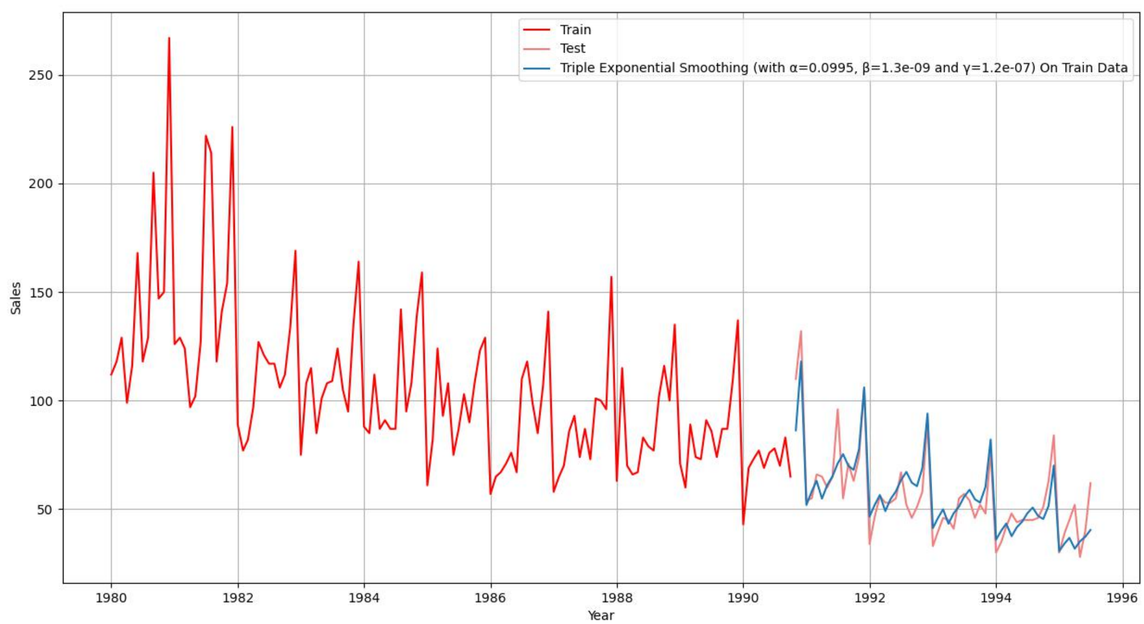
Here, we see that the Double Exponential Smoothing has actually done well when compared to the Simple Exponential Smoothing. This is because of the fact that the Double Exponential Smoothing model has picked up the trend component as well.

Triple Exponential Model

6.4.5 Model Building

Holt-Winters ETS(A,A,A) Model (Additive Error, Additive Trend, Additive Seasonality) model is built on train data and fitted on test data to get the following plot.

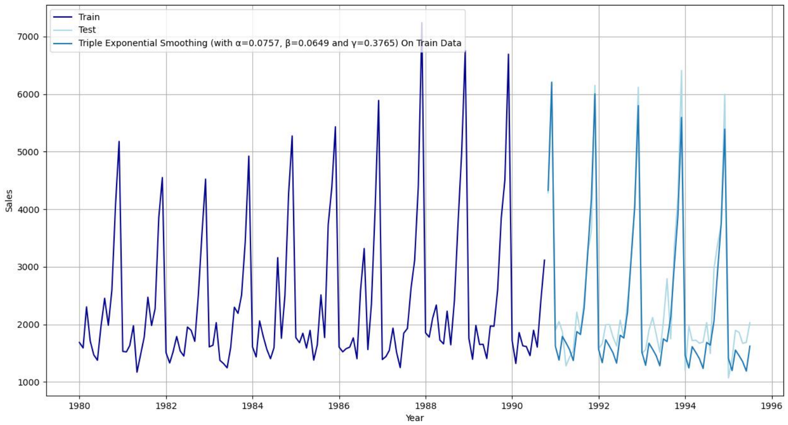
Rose Sales




```
{'smoothing_level': 0.0889188731139181,
 'smoothing_trend': 4.656525272079209e-06,
 'smoothing_seasonal': 0.0,
 'damping_trend': nan,
 'initial_level': 146.86076706887187,
 'initial_trend': -0.5555953323059539,
 'initial_seasons': array([-31.14573949, -18.77116887, -10.7601019 , -21.38589007,
       -12.55757961,  -7.09157308,   2.82817666,   8.93035772,
         4.94016583,   3.0419684 ,  19.65855714,  63.91429684]),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

Figure 27: **Rose sales(Triple Exponential Model and its parameters)**

Sparking Sales




```
{'smoothing_level': 0.07569306568824088,
 'smoothing_trend': 0.03243079354510322,
 'smoothing_seasonal': 0.47913593977130997,
 'damping_trend': nan,
 'initial_level': 2356.5265323809604,
 'initial_trend': -0.7790797250530854,
 'initial_seasons': array([-636.24268485, -722.99133041, -398.6335455 , -473.44377819,
        -808.44186271, -815.36460194, -384.23726708,   72.99614943,
        -237.4576044 ,  272.31934737, 1541.39471434, 2590.08838066]),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```


Figure 28: **Sparkling sales(Triple Exponential Model and its parameters)**

6.4.6 Model Evaluation

The RMSE for the forecast is given below.

	Test RMSE(Rose)	Test RMSE(Sparkling)
Triple Exponential Smoothing (with α=0.0889, β=4.66e-06 and γ=0)	13.976942	NaN
Triple Exponential Smoothing (with α=0.07569, β=0.0324 and γ=0.479)	NaN	366.859156

Triple Exponential Smoothing has performed the best on the test as expected since the data had both trend and seasonality.

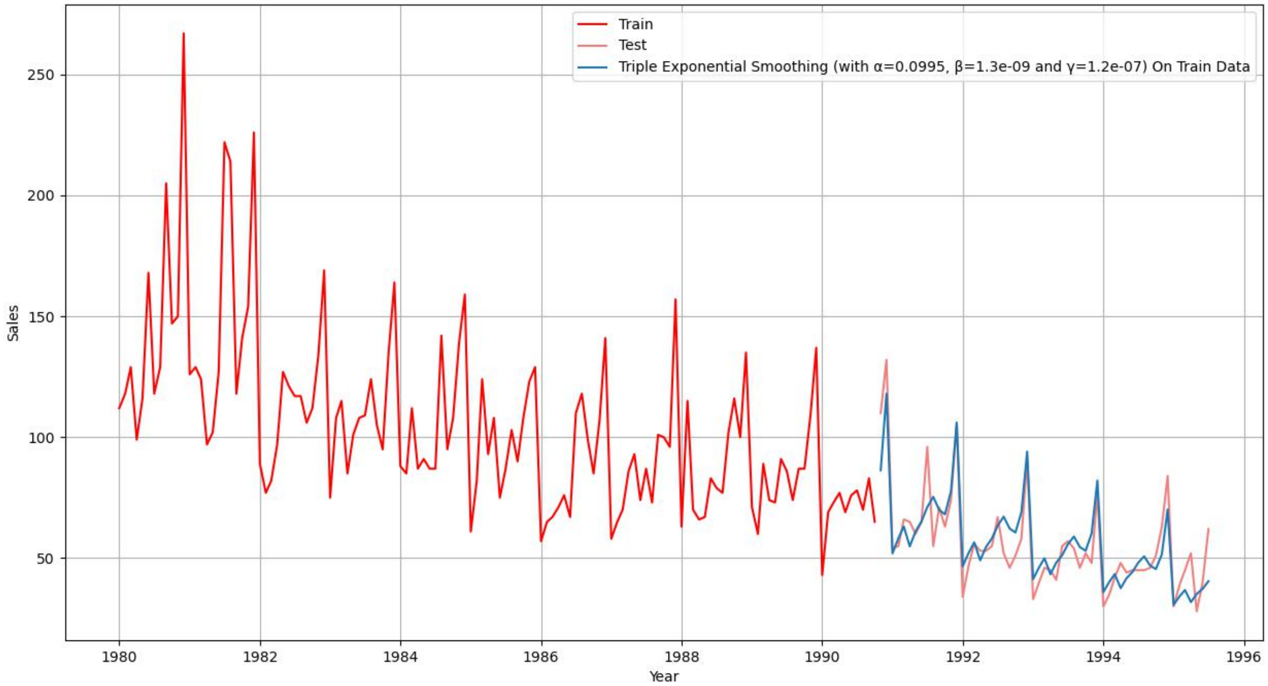
But we see that our triple exponential smoothing is under forecasting. Let us try to tweak some of the parameters in order to get a better forecast on the test set.

Now I have built another model i.e. **Holt-Winters ETS(A,A,M) Model (Additive Error, Additive Trend, Multiplicative Seasonality)**

6.4.7 Model Building

Holt-Winters ETS(A,A,M) model is built on train data and fitted on test data to get the following plot.

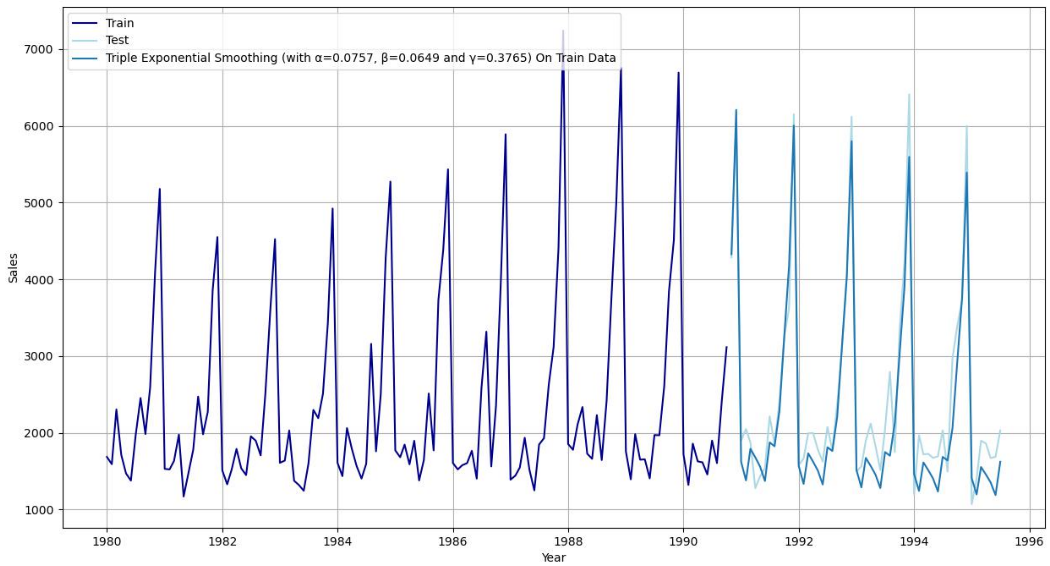
Rose Sales




```
{'smoothing_level': 0.09954161352526007,
 'smoothing_trend': 1.3336303508710234e-09,
 'smoothing_seasonal': 1.2069328449342624e-07,
 'damping_trend': nan,
 'initial_level': 158.17849976224244,
 'initial_trend': -0.6388610685846876,
 'initial_seasons': array([0.69310878, 0.78335434, 0.8565105 , 0.75118137, 0.84323397,
        0.90831655, 0.99998662, 1.06934491, 1.00122429, 0.98484092,
        1.13241501, 1.56136821]),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

Figure 29: **Rose sales(Triple Exponential Model) and its parameters**

Sparking Sales




```
{'smoothing_level': 0.07571432471504627,
 'smoothing_trend': 0.06489794789923221,
 'smoothing_seasonal': 0.3765611795178487,
 'damping_trend': nan,
 'initial_level': 2356.5416847960546,
 'initial_trend': -9.182360270735833,
 'initial_seasons': array([0.71216394, 0.67829895, 0.89649052, 0.79723125, 0.64100433,
        0.63985644, 0.86674058, 1.1133546 , 0.89819179, 1.18511974,
        1.83459596, 2.32779881]),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

Figure 30: **Sparkling sales(Triple Exponential Model and its parameters)**

6.4.8 Model Evaluation

The RMSE for the forecast is given below.

	Test RMSE(Rose)	Test RMSE(Sparkling)
Triple Exponential Smoothing (with α=0.0995, β=1.3e-09 and γ=1.2e-07)	9.334081	NaN
Triple Exponential Smoothing (with α=0.0757, β=0.0649 and γ=0.3765)	NaN	381.655272

Triple Exponential Smoothing has performed the best on the test as expected since the data had both trend and seasonality.

7 Check for Stationarity

Stationarity in Time Series

A time series is considered to be **stationary** when its statistical properties such as the mean, variance, and autocorrelation remain constant over time. Stationarity allows us to model and forecast time series using historical data because the behavior of the series does not change over time. In a stationary series, the autocorrelation at lag k depends only on k, and not on the specific time t. Let X_t denote the value of the time series at time t. The **autocorrelation at lag** k is the correlation between X_t and X_{t-k}.

Checking for Stationarity

To test for stationarity, we use the **Dickey-Fuller test** (specifically, the Augmented Dickey-Fuller or ADF test).

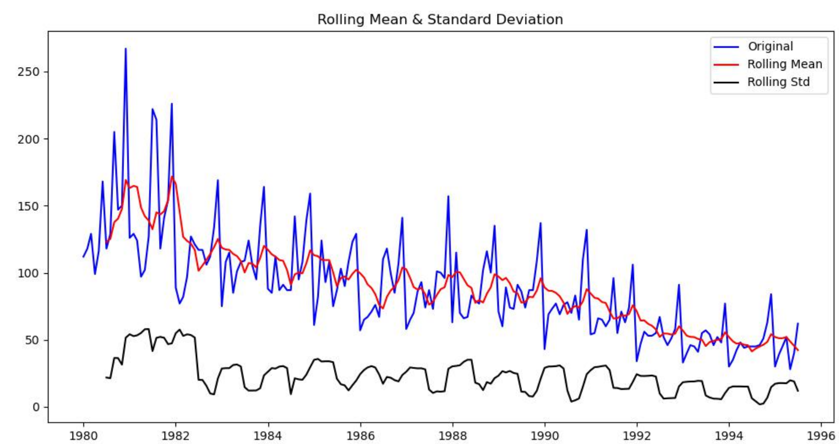
- **Null Hypothesis (H_0)**: The time series is *non-stationary*.
- **Alternative Hypothesis (H_1)**: The time series is *stationary*.

Interpretation:

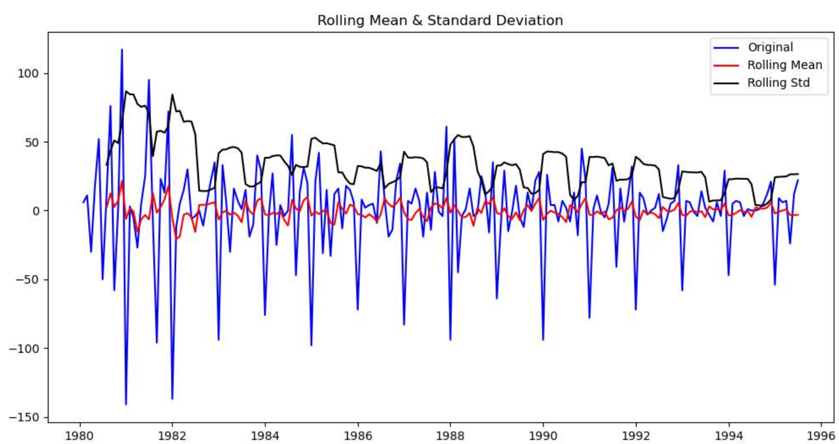
- If the p-value < 0.05: Reject the null hypothesis $\Rightarrow$ The time series is **stationary**.
- If the p-value ≥ 0.05: Fail to reject the null hypothesis $\Rightarrow$ The time series is **non-stationary**.

The test is depicted below. The data is also differentiated once and then checked for stationarity and we find that on single order differencing we are able to get a stationary data.

7.1 Rose Sales



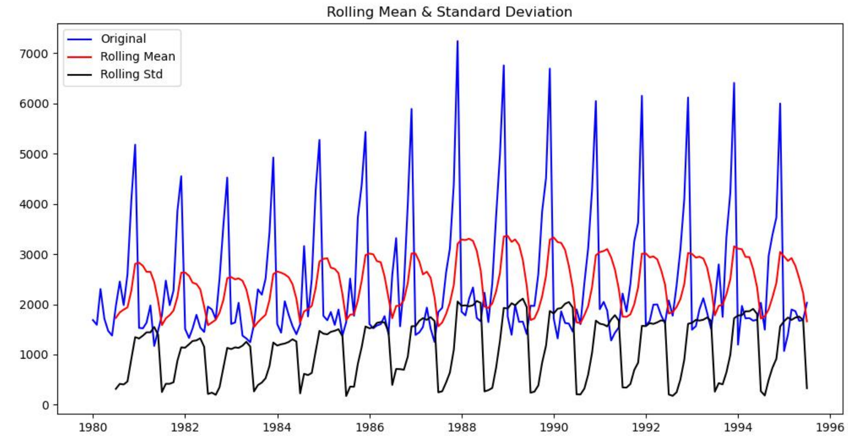

```
Results of Dickey-Fuller Test:
Test Statistic                  -1.874856
p-value                          0.343981
#Lags Used                      13.000000
Number of Observations Used    173.000000
Critical Value (1%)             -3.468726
Critical Value (5%)             -2.878396
Critical Value (10%)            -2.575756
dtype: float64
```



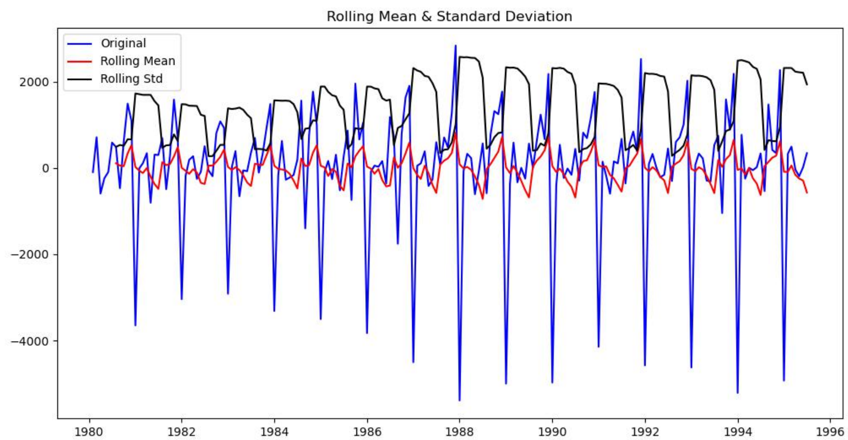

```
Results of Dickey-Fuller Test:
Test Statistic                 -8.044139e+00
p-value                         1.813580e-12
#Lags Used                      1.200000e+01
Number of Observations Used     1.730000e+02
Critical Value (1%)            -3.468726e+00
Critical Value (5%)            -2.878396e+00
Critical Value (10%)           -2.575756e+00
dtype: float64
```

Figure 31: Dickey-Fuller test results for Rose dataset with and without differencing the data

7.2 Sparkling Sales




```
Results of Dickey-Fuller Test:
Test Statistic                  -1.360497
p-value                          0.601061
#Lags Used                      11.000000
Number of Observations Used    175.000000
Critical Value (1%)             -3.468280
Critical Value (5%)             -2.878202
Critical Value (10%)            -2.575653
dtype: float64
```




```
Results of Dickey-Fuller Test:
Test Statistic                 -45.050301
p-value                          0.000000
#Lags Used                      10.000000
Number of Observations Used    175.000000
Critical Value (1%)             -3.468280
Critical Value (5%)             -2.878202
Critical Value (10%)            -2.575653
dtype: float64
```

Figure 32: Dickey-Fuller test results for Sparkling dataset with and without differencing the data

8 Model building of stationary data

8.1 Plotting the Autocorrelation function plots

Using ACF and PACF in Time Series Forecasting

ACF (Autocorrelation Function) and **PACF (Partial Autocorrelation Function)** plots help identify the order of ARIMA models.

- **ACF Plot**: Shows correlation between the time series and its lags. Useful for identifying the MA (Moving Average) order q.
- **PACF Plot**: Shows correlation between the series and its lags after removing intermediate effects. Helps determine the AR (AutoRegressive) order p.

General Guidelines:

- If ACF cuts off after lag q and PACF tails off $\Rightarrow$ MA(q) model.
- If PACF cuts off after lag p and ACF tails off $\Rightarrow$ AR(p) model.

- If both tail off slowly ⇒ Consider ARMA or ARIMA model with differencing.

Note: Bars that lie outside the blue confidence region (typically at 95%) are considered statistically significant.

Rose Sales

8.1.1 Plotting PACF

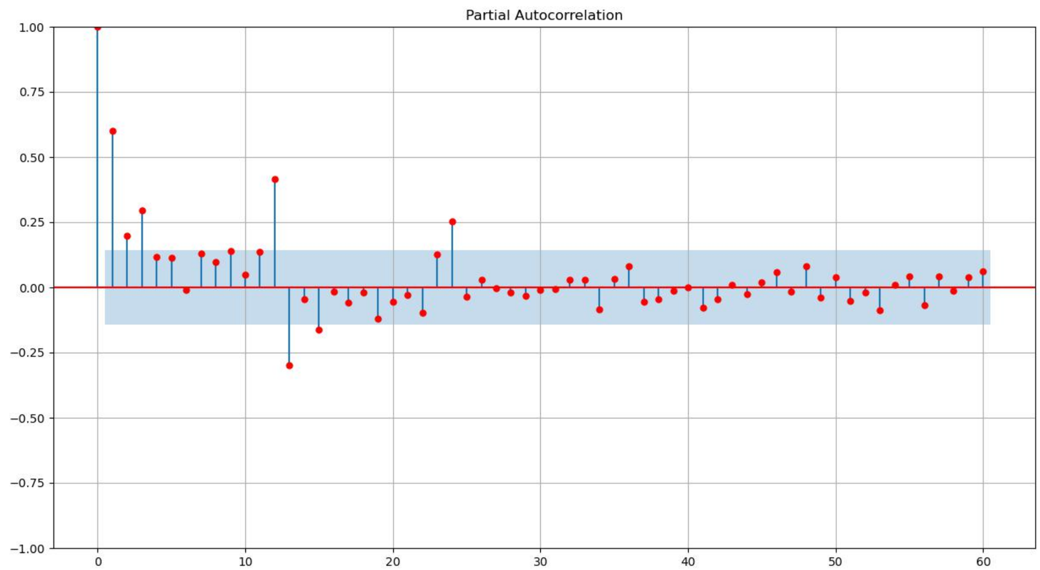


Figure 33: **Rose sales(PACF plot)**

8.1.2 Plotting ACF

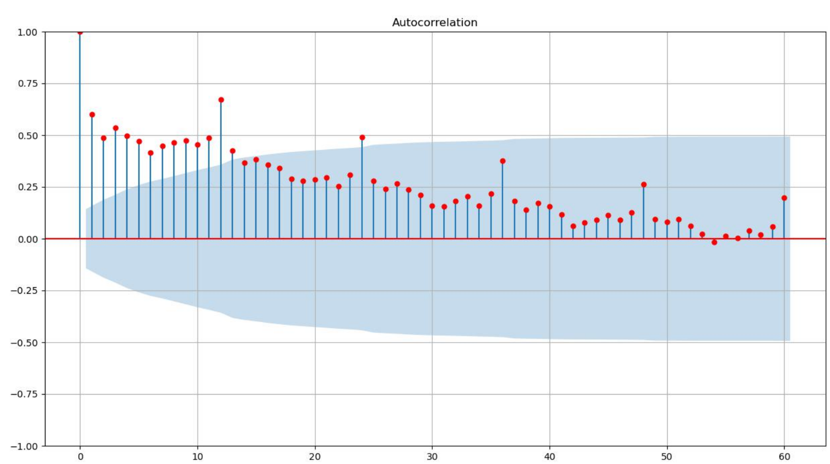


Figure 34: **Rose sales(ACF plot)**

Sparkling Sales

8.1.3 Plotting PACF

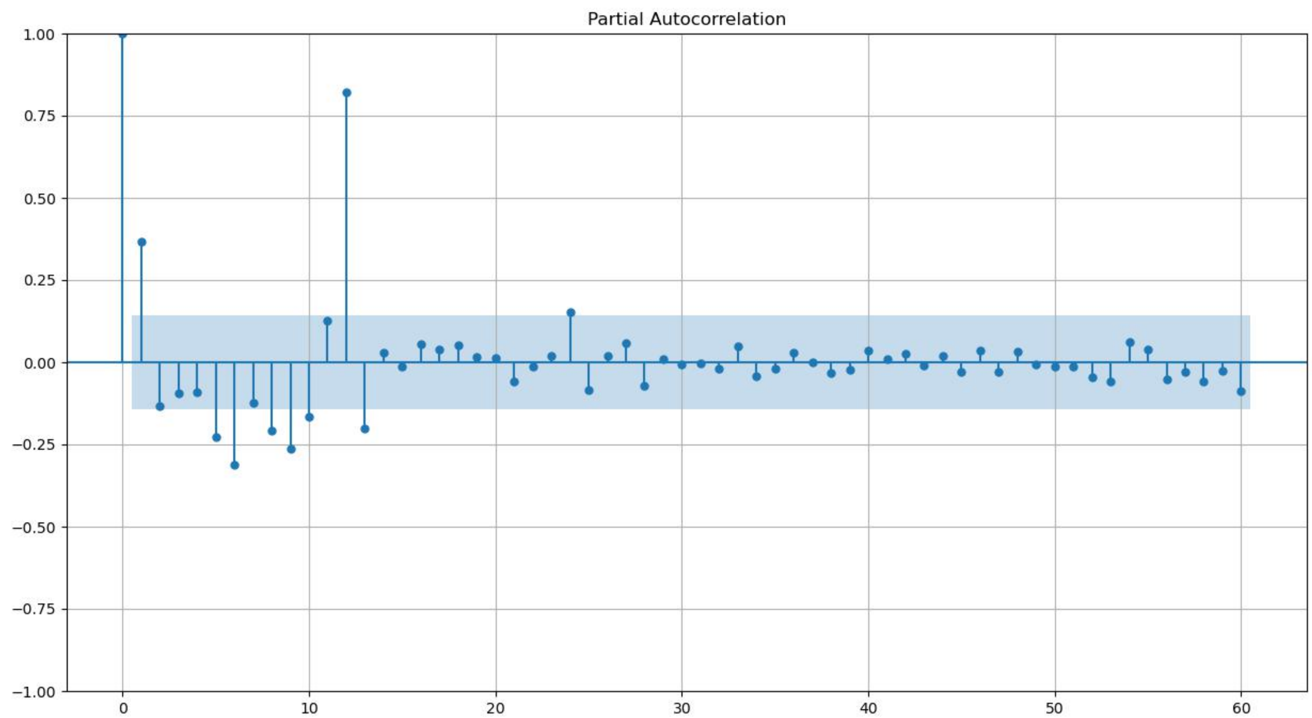


Figure 35: **Sparkling sales(PACF plot)**

8.1.4 Plotting ACF

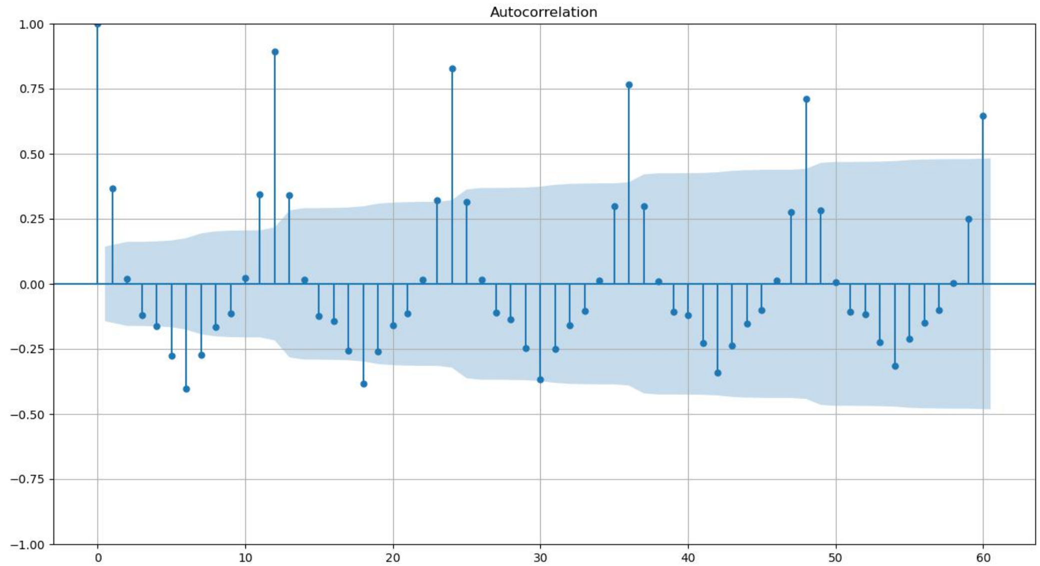


Figure 36: **Sparkling sales(ACF plot)**

We can clearly observe the seasonality from the ACF plot.

Note: The data has some seasonality so ideally we should build a SARIMA model. But for demonstration purposes we are building an ARIMA model both by looking at the minimum AIC criterion and by looking at the ACF and the PACF plots.

8.2 Manual ARIMA Model

p,d,q values are changed and all permutations are tried to get the least AIC value. I have built the best 2 ARIMA models to compare.

```
For Rose
ARIMA(0, 0, 0) - AIC:1306.2886900845197
ARIMA(0, 0, 1) - AIC:1287.0930680691954
ARIMA(0, 0, 2) - AIC:1288.116815603096
ARIMA(0, 1, 0) - AIC:1313.1758613526429
ARIMA(0, 1, 1) - AIC:1261.3274438405808
ARIMA(0, 1, 2) - AIC:1259.2477803151237
ARIMA(1, 0, 0) - AIC:1282.884243262672
ARIMA(1, 0, 1) - AIC:1273.9696715616656
ARIMA(1, 0, 2) - AIC:1272.008960561904
ARIMA(1, 1, 0) - AIC:1297.0772943848615
ARIMA(1, 1, 1) - AIC:1260.0367627036055
ARIMA(1, 1, 2) - AIC:1259.4732049501204
ARIMA(2, 0, 0) - AIC:1283.4621278745708
ARIMA(2, 0, 1) - AIC:1272.7845600980152
ARIMA(2, 0, 2) - AIC:1272.2305289579135
ARIMA(2, 1, 0) - AIC:1278.1352807484318
ARIMA(2, 1, 1) - AIC:1261.0140762916922
ARIMA(2, 1, 2) - AIC:1261.472000656906
```

8.2.1 (Rose Sales)ARIMA model with p=0,d=1 and q=2

The model result is as follows.

SARIMAX Results

Dep. Variable:	Rose	No. Observations:	130
Model:	ARIMA(0, 1, 2)	Log Likelihood	-626.624
Date:	Sun, 06 Apr 2025	AIC	1259.248
Time:	18:58:38	BIC	1267.827
Sample:	01-01-1980	HQIC	1262.734
	- 10-01-1990		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ma.L1	-0.7059	0.072	-9.851	0.000	-0.846	-0.565
ma.L2	-0.1915	0.074	-2.574	0.010	-0.337	-0.046
sigma2	958.5998	86.875	11.034	0.000	788.328	1128.872

Ljung-Box (L1) (Q):	0.15	Jarque-Bera (JB):	45.85
Prob(Q):	0.70	Prob(JB):	0.00
Heteroskedasticity (H):	0.32	Skew:	0.88
Prob(H) (two-sided):	0.00	Kurtosis:	5.34

Figure 37: (Rose Sales)ARIMA model with p=0,d=1 and q=2

All the coefficients are statistically significant.

8.2.2 (Rose Sales)ARIMA model with p=1,d=1 and q=2

The model result is as follows.

SARIMAX Results

Dep. Variable:	Rose	No. Observations:	130
Model:	ARIMA(1, 1, 2)	Log Likelihood	-625.737
Date:	Sun, 06 Apr 2025	AIC	1259.473
Time:	18:58:39	BIC	1270.912
Sample:	01-01-1980	HQIC	1264.121
	- 10-01-1990		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.4649	0.274	-1.698	0.090	-1.002	0.072
ma.L1	-0.2485	0.253	-0.983	0.326	-0.744	0.247
ma.L2	-0.5971	0.208	-2.874	0.004	-1.004	-0.190
sigma2	945.0250	87.810	10.762	0.000	772.921	1117.129

Ljung-Box (L1) (Q):	0.03	Jarque-Bera (JB):	40.04
Prob(Q):	0.86	Prob(JB):	0.00
Heteroskedasticity (H):	0.33	Skew:	0.84
Prob(H) (two-sided):	0.00	Kurtosis:	5.14

Figure 38: (Rose Sales)ARIMA model with p=1,d=1 and q=2

Two of the coefficients are not statistically significant considering 95% confidence.

```
For Sparkling
ARIMA(0, 0, 0) - AIC:2228.4836604091106
ARIMA(0, 0, 1) - AIC:2204.8697988529457
ARIMA(0, 0, 2) - AIC:2206.111207435096
ARIMA(0, 1, 0) - AIC:2232.719438106631
ARIMA(0, 1, 1) - AIC:2217.9392215777407
ARIMA(0, 1, 2) - AIC:2194.0343613616024
ARIMA(1, 0, 0) - AIC:2207.5021008952044
ARIMA(1, 0, 1) - AIC:2206.142158237963
ARIMA(1, 0, 2) - AIC:2207.163048180553
ARIMA(1, 1, 0) - AIC:2231.137663012458
ARIMA(1, 1, 1) - AIC:2196.050085997568
ARIMA(1, 1, 2) - AIC:2194.95965339192
ARIMA(2, 0, 0) - AIC:2204.8807219201435
ARIMA(2, 0, 1) - AIC:2197.084442106501
ARIMA(2, 0, 2) - AIC:2208.1208886931804
ARIMA(2, 1, 0) - AIC:2223.899470277437
ARIMA(2, 1, 1) - AIC:2193.9749624358974
ARIMA(2, 1, 2) - AIC:2178.1097234032827
```

8.2.3 (Sparkling Sales)ARIMA model with p=2,d=1 and q=2

The model result is as follows.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:              Sparkling   No. Observations:                  130
Model:                 ARIMA(2, 1, 2)   Log Likelihood               -1084.055
Date:                Sun, 06 Apr 2025   AIC                           2178.110
Time:                        18:58:41   BIC                           2192.409
Sample:                    01-01-1980   HQIC                          2183.920
                         - 10-01-1990
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          1.3020      0.046     28.542      0.000       1.213       1.391
ar.L2         -0.5360      0.079     -6.763      0.000      -0.691      -0.381
ma.L1         -1.9916      0.109    -18.211      0.000      -2.206      -1.777
ma.L2          0.9998      0.110      9.103      0.000       0.785       1.215
sigma2      1.085e+06   2.03e-07   5.35e+12      0.000    1.08e+06    1.08e+06
===================================================================================
Ljung-Box (L1) (Q):                   0.10   Jarque-Bera (JB):                19.54
Prob(Q):                              0.75   Prob(JB):                         0.00
Heteroskedasticity (H):               2.30   Skew:                             0.71
Prob(H) (two-sided):                  0.01   Kurtosis:                         4.27
===================================================================================
```

Figure 39: (Sparkling Sales)ARIMA model with p=2,d=1 and q=2

All of the coefficients are statistically significant considering 95% confidence.

8.2.4 (Sparkling Sales)ARIMA model with p=2,d=1 and q=1

The model result is as follows.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:              Sparkling   No. Observations:                  130
Model:                 ARIMA(2, 1, 1)   Log Likelihood               -1092.987
Date:                Sun, 06 Apr 2025   AIC                           2193.975
Time:                        18:58:43   BIC                           2205.414
Sample:                    01-01-1980   HQIC                          2198.623
                         - 10-01-1990
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.4862      0.104      4.660      0.000       0.282       0.691
ar.L2         -0.1764      0.190     -0.929      0.353      -0.548       0.196
ma.L1         -0.9999      0.098    -10.225      0.000      -1.192      -0.808
sigma2      1.292e+06   7.62e-08    1.7e+13      0.000    1.29e+06    1.29e+06
===================================================================================
Ljung-Box (L1) (Q):                   0.06   Jarque-Bera (JB):                19.61
Prob(Q):                              0.80   Prob(JB):                         0.00
Heteroskedasticity (H):               2.45   Skew:                             0.67
Prob(H) (two-sided):                  0.00   Kurtosis:                         4.37
===================================================================================
```

Figure 40: (Sparkling Sales)ARIMA model with p=2,d=1 and q=1

one of the coefficients is not statistically significant considering 95% confidence.

	Test RMSE(Rose)	Test RMSE(Sparkling)	Test RMSE(Rose)	Test RMSE(Sparkling)
ARIMA model with p=0,d=1 and q=2	30.923376	NaN	NaN	NaN
ARIMA model with p=2,d=1 and q=2	NaN	1325.16625	NaN	NaN
ARIMA model with p=1,d=1 and q=2	NaN	NaN	30.487641	NaN
ARIMA model with p=2,d=1 and q=1	NaN	NaN	NaN	1359.649838

Figure 41: Comparison of all the ARIMA models

The best ARIMA model for Rose and Sparkling data are shown below.

	Test RMSE(Sparkling)	Test RMSE(Rose)
ARIMA model with p=2,d=1 and q=2	1325.16625	NaN
ARIMA model with p=1,d=1 and q=2	NaN	30.487641

The plots are obtained by forecasting the test data using the best models are attached below.

Rose Sales

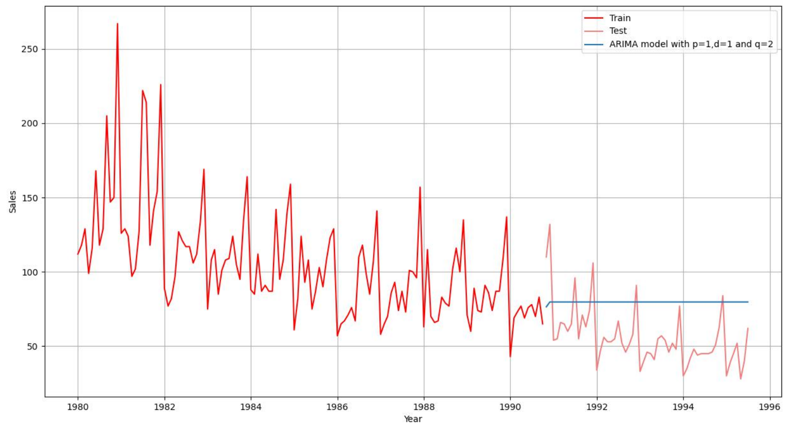


Figure 42: **Rose sales(**ARIMA model with p=1,d=1 and q=2**)**

Sparking Sales

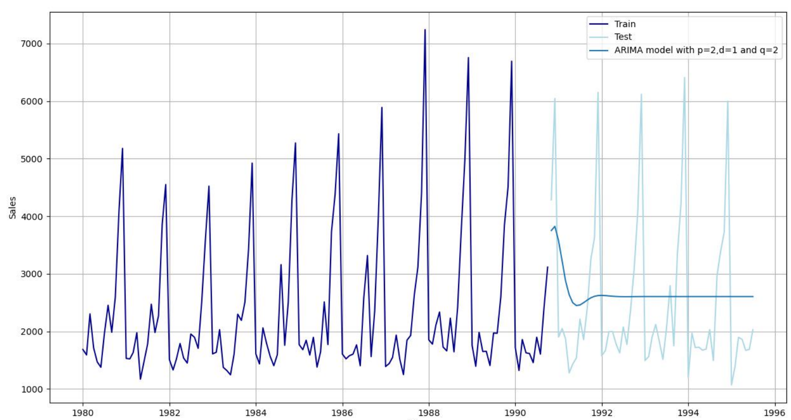


Figure 43: **Sparkling sales(**ARIMA model with p=2,d=1 and q=2 **)**

8.3 Auto ARIMA Model

Auto ARIMA Model is built by using pmdarima library. For Rose data we get the best model to be ARIMA model with p=0,d=1 and q=2 which I have already built. And for Sparkling data the best model comes out to be ARIMA model with p=0,d=0 and q=1.

	Test RMSE(Sparkling)
ARIMA model with p=0,d=0 and q=1	1362.760895

RMSE is more than the best model that I built using ARIMA model with p=2,d=1 and q=2. Hence, I can safely ignore this model.

8.4 Manual SARIMA Model

p,d,q parameters and Seasonal parameters (P,D,Q)S values are changed and all permutations are tried to get the least AIC value. I have built the best SARIMA model to compare.

	parameters	Seasonal parameters	AIC
53	(0, 1, 2)	(2, 1, 2, 12)	759.854456
107	(1, 1, 2)	(2, 1, 2, 12)	761.834085
161	(2, 1, 2)	(2, 1, 2, 12)	763.820849
35	(0, 1, 1)	(2, 1, 2, 12)	767.375258
41	(0, 1, 2)	(0, 1, 2, 12)	768.421506
...	...	...	...
18	(0, 1, 1)	(0, 0, 0, 12)	1242.576606
126	(2, 1, 1)	(0, 0, 0, 12)	1242.720024
108	(2, 1, 0)	(0, 0, 0, 12)	1259.783325
54	(1, 1, 0)	(0, 0, 0, 12)	1287.886350
0	(0, 1, 0)	(0, 0, 0, 12)	1303.984314

162 rows × 3 columns

8.4.1 (Rose Sales)SARIMA model ((0,1,2)(2, 1, 2, 12))

The model result is as follows.

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                                Rose   No. Observations:                  130
Model:             SARIMAX(0, 1, 2)x(2, 1, 2, 12)   Log Likelihood                -372.927
Date:                             Sun, 06 Apr 2025   AIC                            759.854
Time:                                     19:26:30   BIC                            777.353
Sample:                                 01-01-1980   HQIC                           766.911
                                      - 10-01-1990
Covariance Type:                               opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ma.L1         -0.9328      0.191     -4.879      0.000      -1.308      -0.558
ma.L2         -0.0923      0.126     -0.733      0.463      -0.339       0.154
ar.S.L12       0.0367      0.187      0.197      0.844      -0.329       0.402
ar.S.L24      -0.0421      0.030     -1.427      0.154      -0.100       0.016
ma.S.L12      -0.7249      0.294     -2.469      0.014      -1.300      -0.149
ma.S.L24      -0.0677      0.204     -0.331      0.740      -0.468       0.333
sigma2       194.3949     46.225      4.205      0.000     103.795     284.995
===================================================================================
Ljung-Box (L1) (Q):                   0.05   Jarque-Bera (JB):                 4.90
Prob(Q):                              0.82   Prob(JB):                         0.09
Heteroskedasticity (H):               0.93   Skew:                             0.43
Prob(H) (two-sided):                  0.84   Kurtosis:                         3.75
===================================================================================
```

Figure 44: (Rose Sales)SARIMA model ((0,1,2)(2, 1, 2, 12))

Some of the coefficients are statistically insignificant.

	parameters	Seasonal parameters	AIC
95	(1, 1, 2)	(0, 1, 2, 12)	1351.700368
41	(0, 1, 2)	(0, 1, 2, 12)	1352.195894
101	(1, 1, 2)	(1, 1, 2, 12)	1353.484193
149	(2, 1, 2)	(0, 1, 2, 12)	1353.646137
107	(1, 1, 2)	(2, 1, 2, 12)	1353.678153
...	...	...	...
72	(1, 1, 1)	(0, 0, 0, 12)	2165.914890
108	(2, 1, 0)	(0, 0, 0, 12)	2190.833869
18	(0, 1, 1)	(0, 0, 0, 12)	2193.281680
54	(1, 1, 0)	(0, 0, 0, 12)	2214.851626
0	(0, 1, 0)	(0, 0, 0, 12)	2216.418902

162 rows × 3 columns

8.4.2 (Sparkling Sales)SARIMA model ((1,1,2)(0, 1, 2, 12))

The model result is as follows.

```
                                       SARIMAX Results
==========================================================================================
Dep. Variable:                            Sparkling   No. Observations:                130
Model:             SARIMAX(1, 1, 2)x(0, 1, 2, 12)   Log Likelihood                -669.850
Date:                              Sun, 06 Apr 2025   AIC                           1351.700
Time:                                      19:26:37   BIC                           1366.699
Sample:                                  01-01-1980   HQIC                          1357.749
                                       - 10-01-1990
Covariance Type:                                opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.5355      0.225     -2.380      0.017      -0.977      -0.094
ma.L1         -0.2232      0.259     -0.862      0.389      -0.731       0.284
ma.L2         -0.7768      0.162     -4.802      0.000      -1.094      -0.460
ma.S.L12      -0.3970      0.094     -4.231      0.000      -0.581      -0.213
ma.S.L24      -0.0105      0.137     -0.077      0.938      -0.278       0.257
sigma2      1.649e+05   1.47e-06   1.12e+11      0.000    1.65e+05    1.65e+05
===================================================================================
Ljung-Box (L1) (Q):                   0.01   Jarque-Bera (JB):                 8.58
Prob(Q):                              0.92   Prob(JB):                         0.01
Heteroskedasticity (H):               0.83   Skew:                             0.50
Prob(H) (two-sided):                  0.62   Kurtosis:                         4.14
===================================================================================
```

Figure 45: (Sparkling Sales)SARIMA model ((1,1,2)(0, 1, 2, 12))

All of the coefficients are statistically significant considering 95% confidence.

	Test RMSE(Rose)	Test RMSE(Sparkling)
SARIMA model ((0,1,2)(2, 1, 2, 12))	15.426391	NaN
SARIMA model ((1,1,2)(0, 1, 2, 12))	NaN	440.186889

Figure 46: Comparison of all the SARIMA models

The plots are obtained by forecasting the test data using the best models are attached below.

Rose Sales

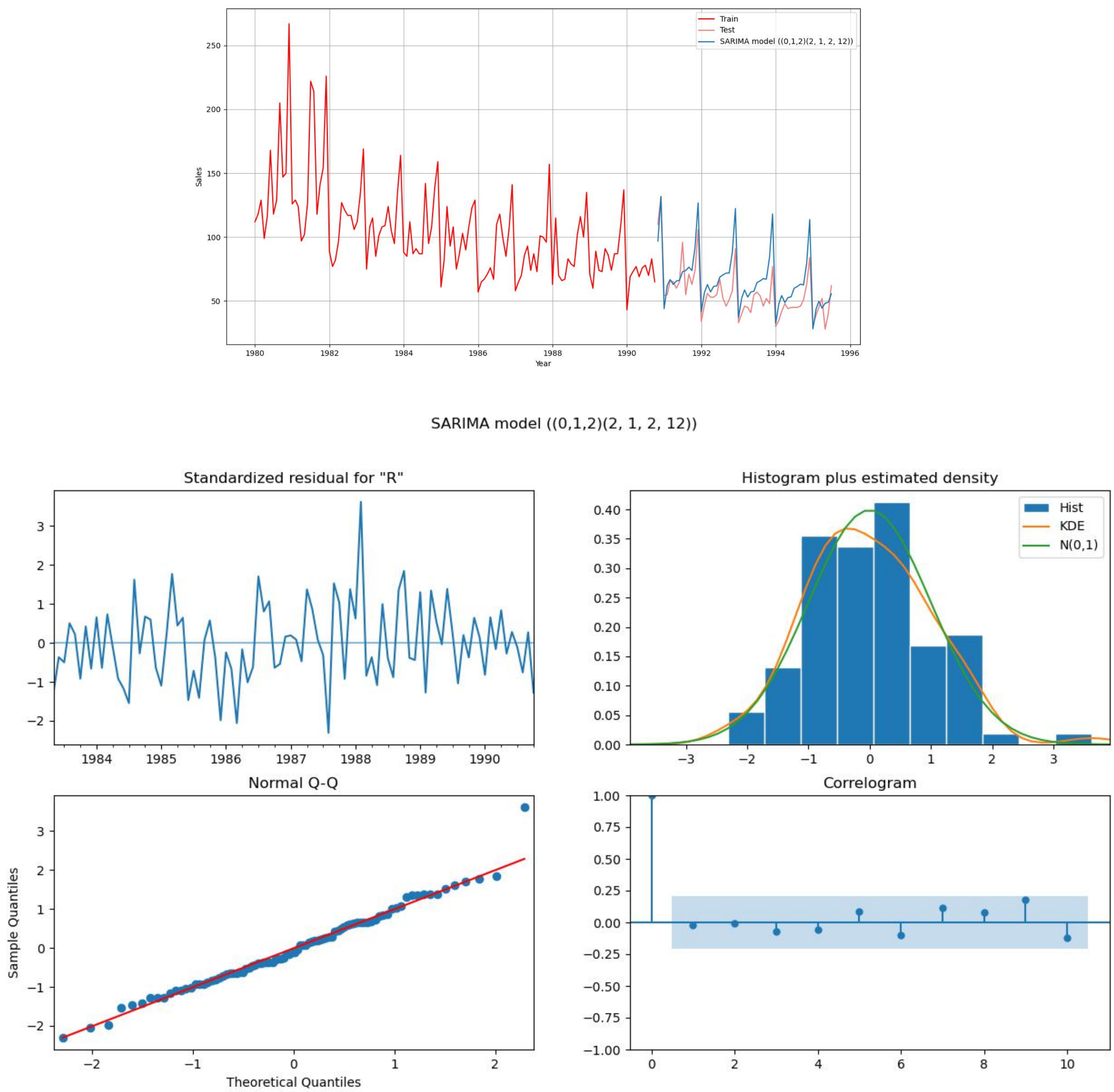
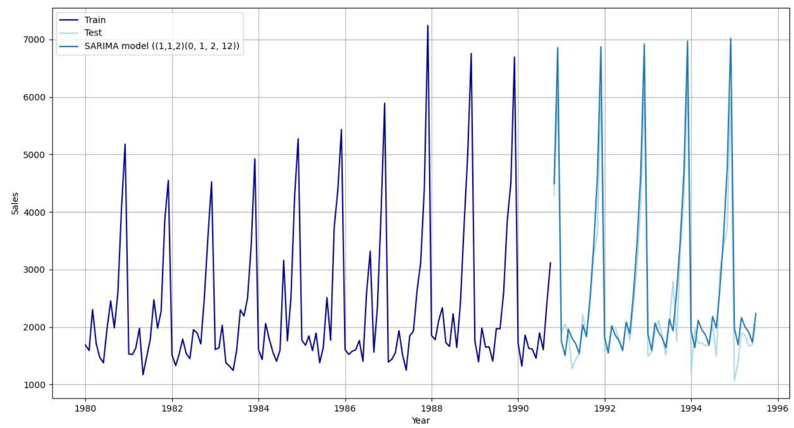


Figure 47: (Rose Sales)SARIMA model ((0,1,2)(2, 1, 2, 12)) and residuals plot

Sparking Sales



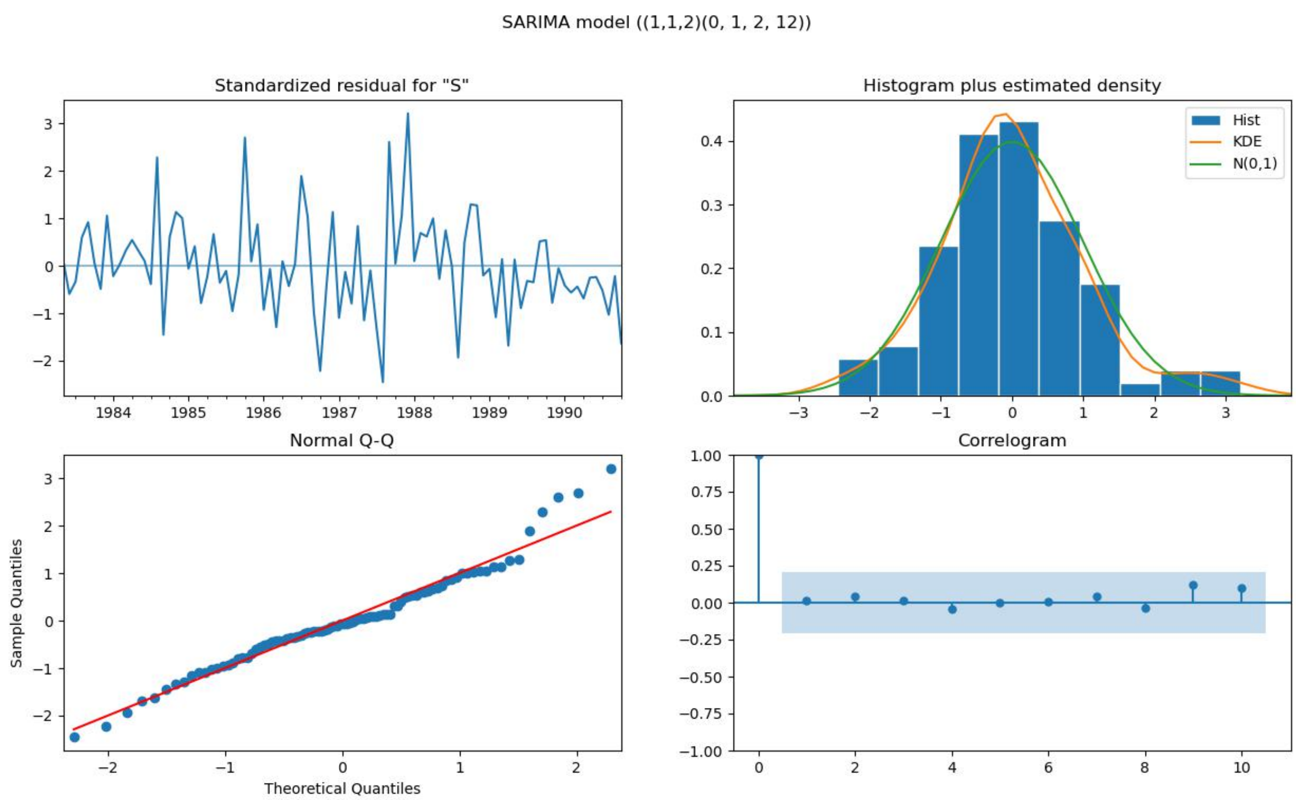


Figure 48: (Sparkling Sales)SARIMA model ((1,1,2)(0, 1, 2, 12)) and residuals plot

8.5 Auto SARIMA Model

Auto SARIMA Model is built by using pmdarima library. For Rose data we get the best model to be SARIMA(5,1,1)(1,0,1)[12] which I have built. (Rose Sales)SARIMA(5,1,1)(1,0,1)[12]

The model result is as follows.

```
Best model:  ARIMA(5,1,1)(1,0,1)[12] intercept
Total fit time: 329.314 seconds
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                                  y   No. Observations:                  130
Model:             SARIMAX(5, 1, 1)x(1, 0, 1, 12)   Log Likelihood                -585.450
Date:                            Sun, 06 Apr 2025   AIC                           1190.900
Time:                                    19:32:33   BIC                           1219.498
Sample:                                01-01-1980   HQIC                          1202.520
                                     - 10-01-1990
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept     -0.0070      0.015     -0.451      0.652      -0.037       0.023
ar.L1          0.2169      0.105      2.067      0.039       0.011       0.423
ar.L2         -0.1995      0.103     -1.935      0.053      -0.402       0.003
ar.L3          0.1316      0.112      1.179      0.239      -0.087       0.350
ar.L4         -0.0923      0.118     -0.782      0.434      -0.324       0.139
ar.L5          0.0696      0.107      0.653      0.514      -0.139       0.278
ma.L1         -0.9348      0.061    -15.436      0.000      -1.054      -0.816
ar.S.L12       0.9858      0.023     43.672      0.000       0.942       1.030
ma.S.L12      -0.8002      0.153     -5.229      0.000      -1.100      -0.500
sigma2       441.2050     63.430      6.956      0.000     316.884     565.526
===================================================================================
Ljung-Box (L1) (Q):                   0.04   Jarque-Bera (JB):                73.70
Prob(Q):                              0.84   Prob(JB):                         0.00
Heteroskedasticity (H):               0.34   Skew:                             0.95
Prob(H) (two-sided):                  0.00   Kurtosis:                         6.17
===================================================================================
```

Figure 49: (Rose Sales)SARIMA(5,1,1)(1,0,1)[12]

Some of the coefficients are statistically insignificant.

8.5.1 (Sparkling Sales)SARIMA(0,0,1)(0,1,1)[12]

The model result is as follows.

```
Best model:  ARIMA(0,0,1)(0,1,1)[12] intercept
Total fit time: 52.544 seconds
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                                  y   No. Observations:                  130
Model:             SARIMAX(0, 0, 1)x(0, 1, 1, 12)   Log Likelihood                -869.820
Date:                            Sun, 06 Apr 2025   AIC                           1747.640
Time:                                    19:33:26   BIC                           1758.723
Sample:                                01-01-1980   HQIC                          1752.140
                                     - 10-01-1990
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept     42.8247     26.717      1.603      0.109      -9.540      95.190
ma.L1          0.1764      0.090      1.953      0.051      -0.001       0.353
ma.S.L12      -0.4656      0.072     -6.491      0.000      -0.606      -0.325
sigma2      1.472e+05   1.39e+04     10.621      0.000     1.2e+05    1.74e+05
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):                48.60
Prob(Q):                              0.95   Prob(JB):                         0.00
Heteroskedasticity (H):               3.37   Skew:                             0.84
Prob(H) (two-sided):                  0.00   Kurtosis:                         5.66
===================================================================================
```

Figure 50: (Sparkling Sales)SARIMA(0,0,1)(0,1,1)[12]

All of the coefficients are statistically significant considering 94% confidence.

	Test RMSE(Rose)	Test RMSE(Sparkling)
AUTO SARIMA model order=(5,1,1),seasonal_order=(1, 0, 1, 12)	16.658045	NaN
AUTO SARIMA model order=(0,0,1),seasonal_order=(0,1,1,12)	NaN	368.352566

Figure 51: Comparison of all the AUTO SARIMA models

9 Comparison of all different models

	Test RMSE(Rose)	Test RMSE(Sparkling)
Linear_Regression	17.356924	1392.438305
Simple_Average	52.431977	1368.746717
Moving_Average	11.801167	811.178937
Simple Exponential Smoothing	30.207858	1363.702251
Double Exponential Smoothing	17.356857	1472.253632
Triple Exponential Smoothing (with α=0.0889, β=4.66e-06 and γ=0)	13.976942	NaN
Triple Exponential Smoothing (with α=0.07569, β=0.0324 and γ=0.479)	NaN	366.859156
Triple Exponential Smoothing (with α=0.0995, β=1.3e-09 and γ=1.2e-07)	9.334081	NaN
Triple Exponential Smoothing (with α=0.0757, β=0.0649 and γ=0.3765)	NaN	381.655272
ARIMA model with p=2,d=1 and q=2	NaN	1325.166250
ARIMA model with p=1,d=1 and q=2	30.487641	NaN
SARIMA model ((0,1,2)(2, 1, 2, 12))	15.426391	NaN
SARIMA model ((1,1,2)(0, 1, 2, 12))	NaN	440.186889
AUTO SARIMA model order=(5,1,1),seasonal_order=(1, 0, 1, 12)	16.658045	NaN
AUTO SARIMA model order=(0,0,1),seasonal_order=(0,1,1,12)	NaN	368.352566

Figure 52: Comparison of all different models

Out of all the models the best model is found to be

Best model for Rose - Triple Exponential Smoothing (with α=0.0995, β=1.3e-09 and γ=1.2e-07)
Best model for Sparkling - Triple Exponential Smoothing (with α=0.07569, β=0.0324 and γ=0.479)

Now, we will take our best model and forecast 12 months into the future with appropriate confidence intervals to see how the predictions look. We have to build our model on the full data for this.

10 Final Forecast

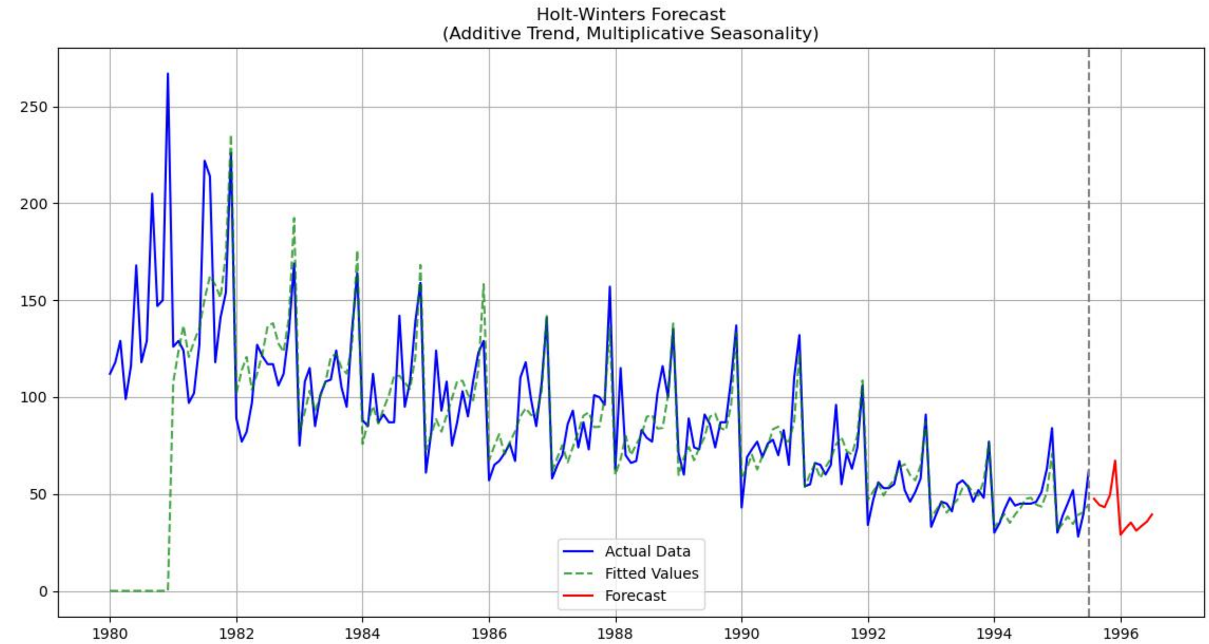


Figure 53: Final forecast plot for Rose

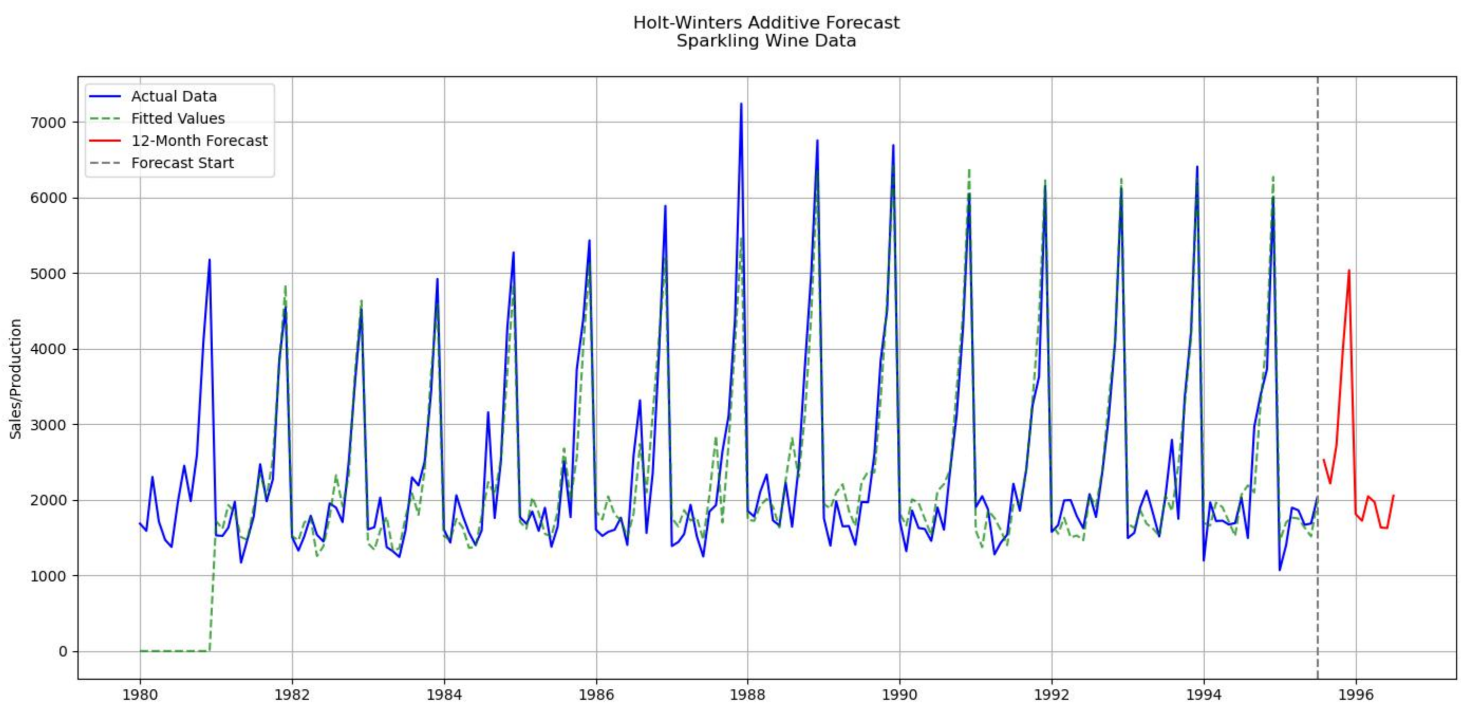


Figure 54: Final forecast plot for Sparkling

11 Actionable Insights and Business Recommendations

Rosé Wine Analysis

- **Trend**: A noticeable decline in overall sales is observed after 1982. Initial years show higher volatility, followed by more stabilized patterns in later years.
- **Seasonality**: Sales display consistent seasonal peaks, with significant surges around November and December, aligning with festive demand.
- **Model Suitability**: The multiplicative model provides a better fit due to the seasonal peaks maintaining a proportional relationship with the overall trend.
- **Business Recommendations**:
  - Launch well-timed seasonal campaigns and festive bundles, especially targeting Q4.
  - Address the long-term downward trend through innovative marketing, rebranding, or introducing limited-edition variants.
  - Investigate patterns during peak early years to extract strategies that can be revived or adapted.
  - Consider diversifying into related beverage segments to counteract declining core demand.

Sparkling Wine Analysis

- **Trend**: The sales trend shows steady growth up to 1995, with some slowdown visible after 1990, though year-end performance remains strong.
- **Seasonality**: Strong and predictable seasonal effects occur yearly, with the highest sales typically in November and December, reflecting strong holiday influence.
- **Model Suitability**: Both additive and multiplicative models fit reasonably well, but the multiplicative model better captures the proportional nature of seasonal fluctuations.
- **Business Recommendations**:
  - Increase inventory and promotional efforts in the second half of the year, especially from July onwards.
  - Capitalize on holiday momentum by introducing special sparkling wine editions or high-margin bundles.
  - Examine strong-performing periods to replicate success factors across markets.
  - Explore export opportunities and upscale positioning in premium product lines.